

UNIVERSIDADE FEDERAL DO PARANÁ

RENATA AMORIM

IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA  
REPRODUÇÃO SEXUADA DO FUNGO ASCOMICETO *Phyllosticta citricarpa*

CURITIBA

2015

RENATA AMORIM

IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA  
REPRODUÇÃO SEXUADA DO FUNGO ASCOMICETO *Phyllosticta citricarpa*

Dissertação de Mestrado apresentada como requisito parcial à obtenção do grau de Mestre em Genética pelo Curso de Pós-Graduação em Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná.

Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Chirlei Glienke

CURITIBA

2015



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO PARANÁ  
SETOR DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE GENÉTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



## PARECER

Os abaixo-assinados, membros da Banca Examinadora da Defesa de Dissertação de Mestrado, a qual se submeteu **RENATA AMORIM**, para obtenção do título de Mestre em Genética pela Universidade Federal do Paraná, no Programa de Pós-Graduação em Genética, são de parecer que se aprove a candidata. Secretária da Coordenação do Programa de Pós-Graduação em Genética do Setor de Ciências Biológicas da Universidade Federal do Paraná.

Curitiba, 22 de setembro de 2015

**Doutora Josiane Aparecida Gomes Figueiredo**  
Universidade Estadual do Paraná – Membro Titular

**Doutora Ana Claudia Bonatto**  
Universidade Federal do Paraná – Membro Titular

**Doutora Vanessa Kava-Cordeiro**  
Universidade Federal do Paraná – Presidente da Banca

Visto

**Profª Drª Lupe Furtado Alle**  
Coordenadora do Programa de Pós-Graduação em Genética

## Dedicatória

Dedico aos meus pais, as únicas pessoas presentes em absolutamente todos os momentos da minha vida.

*“A vontade é a maior de todas as potências.”*

- Léon Denis

## AGRADECIMENTOS

Agradeço a Deus. Não acredito em coincidências, se tenho tantas conquistas e pessoas maravilhosas pelas quais agradecer, é graças a Ele.

Aos meus pais, Susana e Orivan, por todo o apoio, amor e confiança que recebi e recebo. Vocês são os melhores pais que alguém poderia ter.

À minha orientadora, Profa. Dra. Chirlei Glienke, pela oportunidade que me deu em crescer como profissional e como pessoa.

Às Profa. Dra. Vanessa Kava-Cordeiro e Profa. Dra. Lygia Vitória Galli Terasawa, por todo auxílio.

À minha grande e unida família.

Aos meus colegas de laboratório. Ninguém entende tão bem quanto eles os momentos de alegria e desespero que se intercalam com uma velocidade impressionante. Aliandra Medeiros, Francielly Gos, Camila da Costa Senkiv, Paulo José Camargo, Marcos Paulo Rosa, Elvio Henrique Perino, Lorena Carolina Peña.

Lisandra Ferreira Maba, minha companheira de genes *mating type*, que tanto me ensinou, me ajudou e me encorajou quando meu projeto precisou ser mudado no meio do caminho.

Alan Silva, que foi quem me recebeu quando entrei no LabGeM e me ensinou todos os primeiros e mais difíceis passos de quando se entra em um novo desafio.

Juliana Marta e Marcella Cesar, minhas amigas que compartilharam comigo a parte mais difícil: a época da defesa. Obrigada pelos ouvidos e pelo apoio.

Daiani Savi, que me ajudou muito muito muito nas etapas finais desse trabalho. Obrigada!

Aos colegas da Iniciação Científica. Nathália Cordeiro, Kalinka Gonzales, André Servienski, Felipe Gomes, Paulo Henrique Raiser, Nathália Vargas.

Aos técnicos do LabGeM, Rodrigo Aluizio e Ana Paula Chiaverini, por toda a ajuda no laboratório e pela amizade demonstrada ao entrar nas filas quilométricas do RU para almoçar em nossa companhia.

Aos meus amigos, que não permitiram que a insanidade tomasse conta de mim nos momentos mais apurados, e que comemoraram comigo quando a luz no fim do túnel foi vista.

Aos amigos Aderli e Renato Schiochet, que sempre me dizem: a busca pelo horizonte é assim mesmo, não tem fim.

Ao Programa de Pós-Graduação em Genética da Universidade Federal do Paraná.

Ao Departamento de Bioquímica, em especial ao Valter Antônio Baura pelo auxílio técnico na parte do sequenciamento.

A todos do Centro de Citricultura Sylvio Moreira, em especial ao Dr. Marcos Antônio Machado e ao amigo Eduardo Goulin, por toda a ajuda no período em que estive em Cordeirópolis.

Ao pessoal do IAPAR, principalmente Dr. Fabrício Packer Gonçalves, que, apesar do meu projeto ter caminhado em outra direção, sempre me ajudou muito.

À CAPES, por todo o apoio financeiro.

A todos que, de alguma forma, contribuíram com a realização desse trabalho.

## RESUMO

A reprodução sexual em fungos é controlada pelos genes *mating type* presentes no locus *MAT*, que é representado por dois idiomorfos – *MAT1-1* e *MAT1-2*. Em espécies heterotáticas há a presença exclusiva de um dos dois idiomorfos, necessitando o encontro de linhagens com *mating types* opostos para que ocorra a reprodução sexuada. Em contrapartida, espécies homotáticas possuem ambos os idiomorfos em seu genoma, conferindo a esses indivíduos a capacidade de se reproduzir sem necessariamente precisarem de um parceiro compatível. Nesse trabalho nós investigamos a estrutura do locus *MAT* em *Phyllosticta citricarpa*, o agente causal da Mancha Preta dos Citros (MPC). Essa doença causa lesões em frutos e folhas de plantas de citros, onde podem ser encontrados esporos assexuados do fungo (picnidiósporos), enquanto que os esporos sexuados (ascósporos), que são a principal fonte de dispersão da doença em campo, são encontrados em folhas caídas em estado de decomposição no solo. Apesar de intensivo estudo, seu ciclo sexual nunca foi observado em cultura pura *in vitro*. Através da análise das sequências genômicas de dois isolados de *P. citricarpa*, foi identificado o gene da DNA lyase e, como previamente descrito na literatura, os genes *mating type* estavam localizados na região flanqueadora 3' desse gene. Os resultados evidenciam que esse fungo é heterotático devido à presença exclusiva de *MAT1-1* ou *MAT1-2* em linhagens distintas. A fim de caracterizar o locus *MAT*, foram desenhados *primers* para amplificação dessa região. Com base nesses resultados foi demonstrado que *P. citricarpa* possui cópias completas e aparentemente funcionais dos genes *mating type*, contendo os domínios de alta mobilidade (High Mobility Group – HMG) e alpha box, presentes em diferentes isolados. Além dos genes *MAT1-1-1* e *MAT1-2-1*, foram localizados os genes *MAT1-1-4* na região flanqueadora 5' do gene *MAT1-1-1* e o gene *MAT1-2-5*, localizado na região flanqueadora 5' do gene *MAT1-2-1*. Um protocolo de PCR multiplex foi desenvolvido para diferenciar os idiomorfos de *P. citricarpa*, podendo ser utilizado em estudos de distribuição e incidência de genes *mating type*, a fim de determinar a ocorrência de reprodução sexual e facilitar estudos de cruzamento *in vivo* e *in vitro*.

**Palavras-chave:** *Phyllosticta citricarpa*, Mancha Preta dos Citros, reprodução sexual, genes *mating type*.



## ABSTRACT

Sexual reproduction in fungi is controlled by mating type genes located at the *MAT* locus, which is represented by two idiomorphs – *MAT1-1* and *MAT1-2*. In heterothallic species there is the presence of only one of the two idiomorphs, requiring the encounter of opposite mating types strains to occur sexual reproduction. In contrast, homothallic species have both idiomorphs in its genome, giving these individuals the ability to reproduce without necessarily requiring a compatible partner. In this study, we investigated the structure of this locus in the phytopathogenic fungus *Phyllosticta citricarpa*, the epidemiological agent of citrus black spot (CBS) disease. This disease causes lesions on fruits and leaves of citrus plants, where can be found asexual spores of the fungus (picnidiospores), while sexual spores (ascospores), which are the major source of spreading the infection in the field, are found in fallen leaves in state of decomposition. Despite intensive study, its sexual state has never been observed in single-spore culture. Through analysis of the genome sequences of two individual *P. citricarpa* isolates, the sequence of the DNA lyase gene was identified and, as previously reported in the literature, the mating type genes were located in the 3' flanking region of this gene. The results suggested that *P. citricarpa* is heterothallic (self-sterile), owing to the exclusive presence of the *MAT1-1* or *MAT1-2* gene in individual strains. In order to characterize the *MAT* locus, we designed primers to amplify this region. Based on this analysis, *P. citricarpa* was found to have complete and apparently functional copies of *MAT* genes, containing  $\alpha$ -1 and HMG domains, present in different isolates. In addition to *MAT1-2-1* and *MAT1-1-1* genes, the *MAT1-1-4* gene was located in the 5' flanking region of the *MAT1-1-1* gene and the *MAT1-2-5* gene was located in 5' flanking region of the *MAT1-2-1* gene. A multiplex PCR protocol was also developed to differentiate *P. citricarpa* idiomorphs, which can be used in distribution and incidence studies of mating type strains, in order to determine the occurrence of sexual reproduction and to facility crossing studies *in vitro* and *in vivo*.

**Keywords:** *Phyllosticta citricarpa*, Citrus Black Spot, sexual reproduction, *mating type* genes.

## LISTA DE FIGURAS

### CAPÍTULO I

FIGURA 1 - ORGANIZAÇÃO DOS LOCI MAT EM ESPÉCIES HOMOTÁLICAS E HETEROTÁLICAS DO GÊNERO <i>Cochliobolus</i> .....	20
FIGURA 2 - PRIMERS DESENHADOS PARA SEQUENCIAMENTO DA REGIÃO DO LOCUS <i>MAT1</i> DE <i>P. citricarpa</i> .....	28
FIGURA 3 - ORGANIZAÇÃO DOS LOCI <i>MAT1</i> NA ESPÉCIE HETEROTÁLICA <i>P. citricarpa</i> .....	39
FIGURA 4 - PRODUTOS DE AMPLIFICAÇÃO DE PCR MULTIPLEX PARA LINHAGENS DE <i>P. citricarpa</i> <i>MAT1-1</i> (606 PB) E <i>MAT1-2</i> (692 PB).....	40
FIGURA 5 - GRÁFICOS MOSTRANDO AS PROPORÇÕES DOS IDIOMORFOS <i>MAT1-1</i> E <i>MAT1-2</i> NOS POMARES DE CERRO AZUL, DR. ULYSSES E RIBEIRÃO DO SCHEFFER.....	41
FIGURA 6 - ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE AMINOÁCIDOS DO GENE <i>MAT1-1-1</i> .....	42
FIGURA 7 - ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE NUCLEOTÍDEOS DO GENE <i>MAT1-1-1</i> .....	43
FIGURA 8 - ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE AMINOÁCIDOS DO GENE <i>MAT1-2-1</i> .....	44
FIGURA 9 - ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE NUCLEOTÍDEOS DO GENE <i>MAT1-2-1</i> .....	45
FIGURA 10 - ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE NUCLEOTÍDEOS DA ENZIMA DNA Lyase.....	46
FIGURA 11 - ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE AMINOÁCIDOS DA ENZIMA DNA Lyase.....	47

## LISTA DE TABELAS

### CAPÍTULO I

TABELA 1 - LISTA DE <i>PRIMERS</i> E SUAS RESPECTIVAS REGIÕES DE AMPLIFICAÇÃO, SEQUÊNCIAS E CICLOS UTILIZADOS PARA PCR.....	29
TABELA 2 - LINHAGENS DE <i>Phyllosticta citricarpa</i> UTILIZADAS NESSE ESTUDO, MATING TYPE, CÓDIGO DA LINHAGEM, ORIGEM, TECIDO VEGETAL DE ISOLAMENTO, E HOSPEDEIRO.....	32

## SUMÁRIO

1. INTRODUÇÃO .....	13
2. OBJETIVOS .....	16
2.1. Objetivos gerais .....	16
2.2. Objetivos específicos .....	16
3. REVISÃO BIBLIOGRÁFICA.....	17
3.1. <i>Phyllosticta citricarpa</i> e a Mancha Preta dos Citros.....	17
3.2. Reprodução sexuada em ascomicetos.....	18
3.3. PCR Multiplex.....	22
CAPÍTULO I – IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA REPRODUÇÃO SEXUADA DO FUNGO ASCOMICETO <i>Phyllosticta citricarpa</i> ...24	
1. INTRODUÇÃO .....	26
2.2. Identificação do locus <i>MAT</i> a partir das sequências genômicas.....	28
2.2. Desenho de primers específicos para DNA Lyase, <i>MAT1-1-1</i> , <i>MAT1-2-1</i> .....	28
2.3. Extração de DNA e amplificação por PCR .....	29
2.4. Seleção das linhagens de <i>P. citricarpa</i> para sequenciamento .....	30
2.5. Sequenciamento.....	30
2.6. Alinhamento de nucleotídeos, proteínas e análise filogenética .....	30
2.7. PCR Multiplex para identificação dos idiomorfos <i>MAT1-1</i> e <i>MAT1-2</i> .....	31
3. RESULTADOS E DISCUSSÃO .....	34
3.1. Identificação do locus <i>MAT</i> .....	34
3.2. Sequenciamento de linhagens com os idiomorfos <i>MAT1-1</i> e <i>MAT1-2</i> .....	34
3.3. Estrutura do locus <i>MAT1</i> e suas regiões flangeadoras .....	39
3.4. PCR Multiplex para identificação dos idiomorfos <i>MAT1-1</i> e <i>MAT1-2</i> .....	40
3.5. Árvores filogenéticas baseadas em sequências de genes <i>mating type</i> .....	42
CONCLUSÕES.....	51
4. REFERÊNCIAS BIBLIOGRÁFICAS.....	52
5. DISCUSSÃO GERAL .....	58
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	59
ANEXOS.....	68

## 1. INTRODUÇÃO

A produção de laranjas no Brasil é a maior do mundo, correspondendo a cerca de 25% da produção mundial (FAOSTAT, 2015). O país ainda detém 53% da produção mundial do suco dessa fruta, e é responsável por 85% da exportação mundial de suco de laranja (CITRUSBR, 2013). O principal estado produtor de laranja no Brasil é São Paulo, responsável por aproximadamente 70% da produção nacional, seguido por Paraná, Minas Gerais, Bahia e Sergipe (IBGE, 2013). Segundo levantamentos realizados pelo Fundecitrus – Fundo de Defesa da Citricultura, para a safra 2015/2016, estima-se que a produção comercial de laranja para os estados de São Paulo e Minas Gerais seja de 278,9 milhões de caixas de 40,8 kg, e que, dessa produção, 80% será destinada às indústrias processadoras de suco e 20% irão para o mercado *in natura* (FUNDECITRUS, 2015). Devido à sua contribuição econômica para o país, é importante a manutenção da qualidade e rendimento do setor citrícola, que enfrenta grandes desafios relacionados à queda nas exportações, pragas e doenças. Entre essas doenças está a Mancha Preta dos Citros (MPC), causada pelo fungo *Phyllosticta citricarpa*, a qual é caracterizada pelo aparecimento de lesões com borda marrom e um halo amarelo ao redor do centro, onde se encontram pontuações pretas (picnídios). Além da MPC atingir frutos maduros, as lesões podem aparecer também no período de pós-colheita, durante o armazenamento e transporte (FEICHTENBERGER, 1996). Embora as lesões estejam restritas ao flavedo e não afetem a sua qualidade interna, os depreciam visualmente e, conseqüentemente, as mesmas diminuem sua aceitação pelos consumidores no mercado *in natura*. Além de que, em pomares mais gravemente afetados pela doença, há a queda prematura dos frutos (KOTZÉ, 1981) acarretando grandes prejuízos na produção.

O controle da MPC é feito através da aplicação de fungicidas de duas a seis vezes por ano, devido ao longo período de susceptibilidade dos frutos (FUNDECITRUS, 2008). Esse tipo de controle encarece a produção e pode trazer consequências danosas ao ambiente quando não utilizados corretamente. Para criação de alternativas ao controle químico com fungicidas, um maior entendimento da MPC se faz necessário através de estudos sobre a epidemiologia da doença e interações entre a planta e o patógeno.

Acredita-se que na epidemiologia da MPC, os ascósporos tenham um papel determinante na dispersão da doença. Embora Timossi e colaboradores (2003) tenham descrito a produção de ascósporos em laboratório, o ciclo sexual em *P. citricarpa* ainda não está completamente esclarecido e tal produção não foi repetida com sucesso até então. Uma possível explicação seria o heterotalismo de *P. citricarpa*, fazendo-se necessária a inoculação conjunta de linhagens sexualmente compatíveis. Em 2002, Baayen e colaboradores já haviam sugerido que *P. citricarpa* é heterotática ao afirmar que não foi observada a produção de estruturas de reprodução sexuada em laboratório, enquanto que a espécie intimamente relacionada *P. capitalensis* apresentou produção de pseudotécios e, portanto, é evidentemente homotática. Os Ascomicetos são caracterizados pela formação dos ascósporos dentro de uma estrutura denominada asco e são agrupados em classes de acordo com o tipo de corpo de frutificação e forma de liberação dos ascósporos (PÖGGELER *et al.*, 2006).

A reprodução sexual em Ascomicetos é controlada por um único *mating type* locus denominado *MAT1*, que é representado por dois idiomorfos conhecidos como *MAT1-1*, que codifica fatores de transcrição para proteína de domínio alfa, e *MAT1-2*, que codifica fatores de transcrição para proteína de domínio de alta mobilidade (HMG) (TURGEON e YODER, 2000; ARZANLOU, *et al.*, 2010). Desta forma, a caracterização e descrição dos idiomorfos *mating type* visa contribuir com o conhecimento sobre o ciclo sexual de *P. citricarpa*, espécie essa de importância agrícola e contribuir com estudos de reprodução *in vitro*, através da produção de ascósporos e sua utilização em experimentos que necessitem da reprodução de sintomas em laboratório. Embora já existam estudos com outros patógenos relacionando os genes APN2 (DNA Lyase) e SLA2 (controle de montagem do citoesqueleto) como regiões flanqueadoras comuns dos idiomorfos *MAT1-1* e *MAT1-2* (TURGEON *et al.*, 1993; METZENBERG e GLASS, 1990) e a função de regular o ciclo sexual em fungos, estes idiomorfos não apresentam homologia nas sequências de nucleotídeos entre diferentes espécies (KLIX *et al.*, 2010). Compreender os fatores que controlam os modos de reprodução de um organismo é importante porque padrões de herança afetam os principais processos ecológicos e evolutivos (OTTO, 2009), bem como auxiliam no entendimento do ciclo da doença através do estudo da evolução e controle de patógenos (ARZANLOU *et al.*, 2010; TSUI *et al.*, 2013; BIHON *et al.*, 2014).

Neste trabalho, além da identificação e sequenciamento do locus *MAT*, foi desenvolvido um protocolo de PCR Multiplex para facilitar estudos de distribuição e/ou incidência de diferentes *mating types* em determinada região, podendo assim prever o risco de ocorrência de reprodução sexual e disseminação da doença, além de facilitar estudos de cruzamento e testes de indução de sintomas *in vitro*.

## **2. OBJETIVOS**

### **2.1. Objetivos gerais**

Contribuir para o conhecimento do ciclo sexual do fungo *Phyllosticta citricarpa* através do estudo do locus *MAT1*, responsável pela regulação sexual e produção de ascósporos, cuja estrutura é a principal forma de dispersão da Mancha Preta dos Citros.

### **2.2. Objetivos específicos**

- Identificar e sequenciar o locus *MAT1* em linhagens de *P. citricarpa*.
- Caracterizar e descrever os idiomorfos *MAT1-1* e *MAT1-2* em linhagens de *P. citricarpa*.
- Facilitar estudos de reprodução *in vitro*.
- Determinar se os idiomorfos *MAT1* encontram-se em proporções semelhantes em campo.
- Desenvolver PCR Multiplex específico para os idiomorfos *MAT1-1* e *MAT1-2* do locus *MAT1* em *P. citricarpa*.
- Correlacionar filogeneticamente o locus *MAT1* entre espécies pertencentes à classe dos Dotideomicetos.



### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1. *Phyllosticta citricarpa* e a Mancha Preta dos Citros

A Mancha Preta dos Citros (MPC) é causada pelo fungo *Phyllosticta citricarpa* (McAlpine) Aa (1973), cuja fase sexuada foi denominada como *Guignardia citricarpa* Kiely (1948). Em 2011, após o Simpósio “The Amsterdam Declaration on Fungal Nomenclature”, ocorrido em Amsterdã com o propósito de criar um sistema de nomenclatura único para todos os fungos, ficou decidido que a prioridade deve ser dada ao nome cujo gênero foi antes descrito, entrando em vigor a partir de 1º de janeiro de 2013. *Guignardia* foi descrito em 1892 e *Phyllosticta* em 1818, portanto o nome oficialmente em uso é *P. citricarpa* (WULANDARI *et al.*, 2009; HAWKSWORTH *et al.*, 2011).

*P. citricarpa* pertence ao filo Ascomycota, classe Dotideomycetes, ordem Botryosphaerales e, a partir dos estudos filogenéticos, morfológicos e ecológicos realizados por Slippers e colaboradores (2013) e Wikee e colaboradores (2013), *P. citricarpa*, bem como todas as espécies do gênero *Phyllosticta*, foi classificada como pertencente à família *Phyllostictaceae* e não mais à família *Botryosphaeriaceae*.

As espécies do gênero *Phyllosticta* são, em sua maioria, fitopatógenos de uma grande variedade de hospedeiros (GLIENKE *et al.*, 2011; WANG *et al.*, 2012), existindo cerca de 3405 espécies listadas para o gênero *Phyllosticta* no MycoBank (<http://www.mycobank.org/>; acessado Junho 2015), 3215 espécies listadas no Index Fungorum (<http://www.indexfungorum.org/>; acessado Junho 2015), e 32 espécies do gênero relacionadas à espécies de citros, sendo elas patogênicas, endofíticas e/ou sapróbias segundo o USDA Fungal Database (<http://nt.ars-grin.gov/fungaldatabases/>; acessado Junho 2015).

A doença foi descrita pela primeira vez na Austrália em 1895 (SUTTON e WATERSON, 1966), e desde então vem causando prejuízos à cultura dos citros em todos os continentes, exceto aos países da Europa onde a doença não foi detectada, sendo considerada uma doença quarentenária A1. As cargas provenientes de países onde a doença já foi detectada enfrentam barreiras fitossanitárias que impedem a disseminação da doença nesse continente (EFSA, 2014).

O ciclo da MPC começa com a produção de ascósporos em folhas caídas em processo de decomposição (REIS *et al.*, 2006). Os ascósporos são esporos sexuais

liberados em condições de umidade e temperatura ideais, e então dispersos pelo vento e responsáveis pela disseminação da doença a outras plantas de citros (HUANG, 1972; REIS, 2006). Após instalada, a doença pode permanecer latente sem o aparecimento de sintomas por anos (McONIE, 1964a). Em situações ambientais propícias, o fungo causa lesões nas cascas dos frutos onde são produzidos os picnídios, esporos assexuados responsáveis pela dispersão da doença a curtas distâncias, geralmente pela água da chuva, e que agrava a infecção na própria planta (SPÓSITO *et al.*, 2008). A MPC é responsável por grandes perdas em frutos de laranja doce, tanto em frutos nos pomares quanto na fase pós-colheita (SUTTON e WATERSTON, 1966).

Quase todas as variedades de laranja doce (*Citrus sinensis* (L.) Osbeck) e limões (*C. limon* (L.) Burm.) de interesse comercial são suscetíveis à doença, bem como outras variedades de *Citrus* spp. Como exceções, a laranja ácida (*C. aurantium* L.) e seus híbridos que são resistentes à doença (KOTZÉ, 1981) e a lima ácida Tahiti (*C. latifolia*), que embora o fungo possa ser isolado de suas folhas e frutos, não é sensível à doença (BALDASSARI *et al.*; 2008).

### **3.2. Reprodução sexuada em ascomicetos**

Reprodução sexual é uma característica que confere vantagens ecológicas aos organismos, dando origem a uma grande diversidade de genomas com maior habilidade de adaptação a mudanças ambientais quando comparados a populações de reprodução assexuada (KLIX *et al.*, 2010; HEITMAN *et al.*, 2013).

Os fungos possuem diferentes tipos de reprodução sexual, como encontro de sexos opostos, auto-fertilidade, ciclo parassexual ou ainda a presença de múltiplos sexos (KRONSTAD, 2007). O comportamento sexual em fungos pode ser melhor explicado através da regulação pelo locus *mating type* (*MAT1*) que codifica fatores de transcrição responsáveis pela reprodução sexual, além de componentes de sinalização como feromônios, receptores de feromônios, proteínas quinases e outros fatores que controlam o desenvolvimento sexual, como foi descrito em ascomicetos do gênero *Aspergillus* (COPPIN *et al.*, 1997; KRONSTAD, 2007). A grande maioria das espécies pertencentes ao filo Ascomicota contém um único locus *mating type* (*MAT1*) responsável pela sua identidade sexual e pela fertilidade, visto que já foi observado que a supressão desses genes resulta em ausência de ascósporos (PAOLETTI *et al.*, 2007).

Existem espécies em que, para que haja compatibilidade e consigam realizar a reprodução sexuada, deve haver o encontro de dois indivíduos de *mating types* opostos (heterotalismo). Cada linhagem deve possuir um diferente idiomorfo no locus *MAT1*, sendo eles *MAT1-1* que codifica uma proteína regulatória de domínio alpha-box e *MAT1-2* que codifica proteína regulatória de domínio High Mobility Group (HMG) (TURGEON e YODER, 2000). São chamados idiomorfos porque não têm homologia de sequências significativa e codificam para diferentes fatores de transcrição, não podendo assim ser chamados de alelos (YUN *et al.*, 1999). O uso do termo idiomorfo fica restrito a descrever o locus apenas em espécies heterotáticas, sendo inapropriado para espécies homotáticas (DEBUCHY *et al.*, 2010).

Groenewald e colaboradores (2007) desenharam *primers* degenerados capazes de amplificar regiões de ambos os idiomorfos *MAT1-1* e *MAT1-2* dos fungos morfologicamente relacionados *Dothistroma septosporum* e *Dothistroma pini*, responsáveis pela doença Queima de Acículas ou Queima de Dotistroma em pinheiros. Tais *primers* foram capazes de distinguir os dois fungos e os diferentes *mating types* dessas espécies. Já Dyer e colaboradores (2001), não obtiveram o mesmo sucesso com as espécies heterotáticas *Tapesia yallundae* e *Tapesia acuformis*, onde os *primers* desenhados para os genes *MAT* de *T. yallundae* não foram capazes de amplificar os mesmos genes de *T. acuformis*. Li e colaboradores (2013) ainda identificaram elementos transponíveis nos loci *MAT* de algumas linhagens de *Blastomyces dermatidis*, tornando-os excepcionalmente grandes quando comparados aos locus *MAT* de outros fungos dimórficos e dificultando ainda mais sua identificação. O trabalho de Maba (2014) sugere que *P. citricarpa* é uma espécie heterotática, pois foram identificadas linhagens contendo genes *MAT1-1-1* ou *MAT1-2-1*.

Existem espécies que por muito tempo foram consideradas assexuadas, e através de análises moleculares de *mating type* foi demonstrado que possuem a maquinaria necessária para realizar o ciclo sexual por reprodução heterotática. Como por exemplo no caso da espécie de grande valor biotecnológico *Penicillium chrysogenum* (BÖHM *et al.*, 2012) que mesmo sendo estudado a mais de 80 anos ainda não tinha seu ciclo sexual descrito. Outro exemplo é o patógeno oportunista *Aspergillus fumigatus* que teve seu locus *MAT* identificado como heterotático por Paoletti e colaboradores (2005), após analisar 290 linhagens e revelar

a presença dos idiomorfos *MAT1-1* e *MAT1-2* em iguais proporções. Os autores afirmam que os dados expostos nesse trabalho indicam que *A. fumigatus* possui uma história evolutiva de recombinação sexual recente e, portanto, potencial para realizar reprodução sexuada. Maba (2014) e Bihon e colaboradores (2014) demonstraram a dificuldade de observar em laboratório a formação de corpos de frutificação em *P. citricarpa* e em *D. pinea*, respectivamente, não obtendo sucesso após inúmeros cruzamentos *in vitro*. Uma possível explicação para isso é de que condições específicas necessárias para a realização do ciclo sexual não tenham sido atingidas.

Os genes *mating type* são também encontrados em espécies homotáticas de ascomicetos, nas quais o micélio é auto fértil e não há necessariamente fecundação cruzada nem tipos de pareamento (AZEVEDO e ESPOSITO, 2010). Algumas dessas espécies possuem os dois genes *MAT* (*MAT1-1-1* e *MAT1-2-1*) presentes no genoma, e podem estar organizados de várias formas: fusionados, ligados em uma mesma região ou em loci separados (YUN *et al.*, 1999; PAOLETTI *et al.*, 2007; DEBUCHY *et al.*, 2010). Homotalismo pode servir como garantia reprodutiva, mantendo genótipos bem adaptados em áreas onde parceiros compatíveis sejam escassos (PAOLETTI *et al.*, 2007). A comparação entre algumas possíveis organizações do locus *MAT* pode ser observada na Figura 1, que mostra espécies do gênero *Cochliobolus* que foram estudadas por Yun e colaboradores (1999) a fim de tentar esclarecer se o homotalismo é derivado do heterotalismo ou vice-versa. Como o locus *MAT* é responsável por controlar o processo reprodutivo, a comparação entre suas sequências pode mostrar mudanças nesse processo ao longo do tempo. Além dos exemplos mostrados por esse autor, existem outras possíveis organizações do locus *MAT* descritas na literatura. Zaffarano e colaboradores (2010) descreveram a estrutura do locus *MAT* do complexo de espécies pertencente à *Phialocheopala fortinii* s.l. – *Acephala applanata*, demonstrando a presença do gene *MAT1-1-3* em algumas das espécies estudadas e até mesmo a presença de um elemento transponível em *Acephala applanata*. Yokoyama e colaboradores (2006) estudaram o locus *MAT* de fungos da família *Clavicipitaceae*, pertencente à classe Sordariomicetos. Esses fungos geralmente possuem 3 genes *mating type* no locus *MAT1-1* – *MAT1-1-1*, *MAT1-1-2* e *MAT1-1-3* – porém, em *Neotyphodium unconatum*, o gene *MAT1-1-3* mostrou-se interrompido por um códon de parada, parecendo ter se tornado um pseudogene.

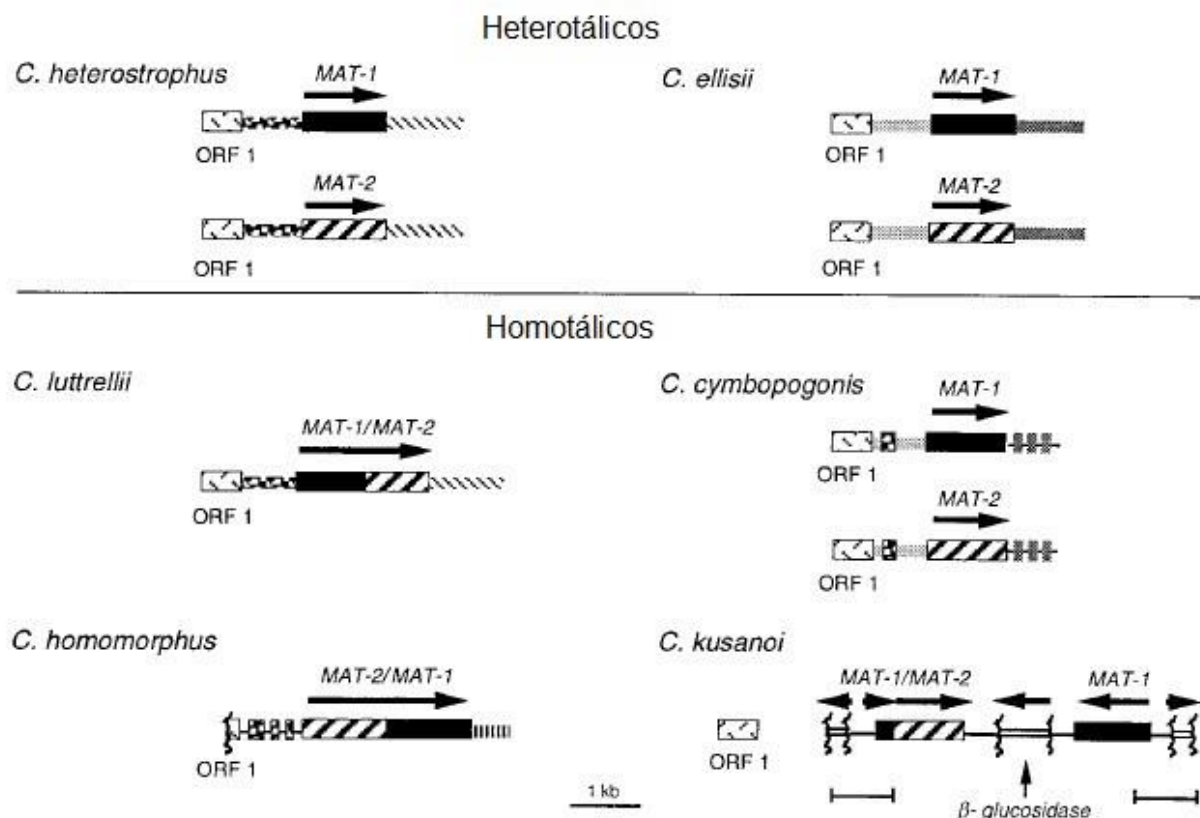


FIGURA 1: ORGANIZAÇÃO DOS LOCI MAT EM ESPÉCIES HOMOTÁLICAS E HETEROTÁLICAS DO GÊNERO *Cochliobolus*.

FONTE: YUN *et al.*, 1999.

Nota: O arranjo para as espécies heterotáticas era basicamente o mesmo. Já para espécies homotáticas cada organização pode ser única, estando os genes fusionados, ligados em uma mesma região ou em loci separados.

Um exemplo da peculiaridade na organização do locus *MAT* em espécies homotáticas é o caso encontrado por Chitrampalam e colaboradores (2013) em linhagens de *Sclerotinia sclerotinum*, que é um fungo fitopatogênico descrito como homotáticos por possuir os dois genes *MAT* fusionados. Porém, em algumas linhagens estudadas foi encontrada uma região de 3,6 kb de inversão a cada geração meiótica, resultando na inversão dos genes *MAT1-2-1* e *MAT1-2-4* e ainda truncando o gene *MAT1-1-1* na extremidade 3', afetando a expressão gênica entre linhagens com a inversão e sem a inversão. Não se sabe o que desencadeia esse processo, mas em células sanguíneas humanas esses eventos são causados por recombinação homóloga não alélica. Esse tipo de recombinação também já foi descrito em

*Saccharomyces cerevisiae*, resultando em rearranjos do seu genoma (HOANG *et al.*, 2010).

Conhecer os modos de reprodução e os mecanismos envolvidos no ciclo sexual de espécies de interesse agrônomo, bem como conseguir realizar a reprodução sexuada em laboratório são importantes ferramentas para a análise genética desses fungos e da sua correlação com os hospedeiros.

### **3.3. PCR Multiplex**

Reação em cadeia polimerase (PCR) multiplex difere da PCR comum pelo fato de que dois ou mais loci são amplificados simultaneamente na mesma reação (HENEGARIU *et al.*, 1997; ELNIFRO *et al.*, 2000), economizando tempo e reagentes. Essa tem sido uma tecnologia muito utilizada na área de diagnósticos moleculares desde a década de 90 (WOLCOTT, 1992), e desde então são várias as áreas de sua aplicabilidade, dentre elas estão a análise de mutações (TALY *et al.*, 2013), detecção de polimorfismos de nucleotídeo único, inserções e deleções (ZHENG *et al.*, 2014). Na área das doenças infecciosas, se mostrou um bom método para identificação de bactérias, fungos e outros parasitas (HENDOLIN *et al.*, 1997; CÔTÉ *et al.*, 2004; TANIUCHI *et al.*, 2011).

PCR Mutiplex é um método muito utilizado na identificação de fungos de importância agrônoma. Demeke e colaboradores (2005) utilizaram esse teste na identificação de nove espécies de fungo do gênero *Fusarium*, causadores de uma das mais importantes doenças fúngicas de cereais ao redor do mundo, a fusariose do trigo. Chen e colaboradores (2008) identificaram duas diferentes espécies de fungos causadores de oídio em campos de girassol no Taiwan, *Podosphaera xanthii* e *Golovinomyces cichoracearum*, através desse método. Fu e colaboradores (2014) utilizaram PCR Multiplex na identificação do fungo *Villosiclava virens* como homotático, visto que esse apresentou amplificações referentes aos genes *MAT1-1-1* e *MAT1-2-1* nos resultados observados em gel de agarose.

PCR Multiplex pode também ser utilizada para detectar diferentes genes de um mesmo organismo. Poirel e colaboradores (2011) utilizaram essa técnica na identificação de genes codificadores de cabapenemase pertencentes a diferentes classes, e ainda para identificação de genes *mating type*, como Rubini e colaboradores

(2010) que identificaram os diferentes genes *MAT* de *Tuber melanosporum* através de PCR Multiplex.

## **CAPÍTULO I – IDENTIFICAÇÃO E CARACTERIZAÇÃO DE GENES ENVOLVIDOS NA REPRODUÇÃO SEXUADA DO FUNGO ASCOMICETO *Phyllosticta citricarpa***

### **Resumo**

A reprodução sexual em fungos é controlada pelos genes *mating type*, presentes no locus *MAT1*. Nesse trabalho, nós investigamos a estrutura desse locus no fungo fitopatogênico *Phyllosticta citricarpa*, agente causal da Mancha Preta dos Citros. Como em outros Ascomicetos heterotáticos, os idiomorfos presentes no locus *MAT* dessa espécie – *MAT1-1* e *MAT1-2* – codificam para proteínas de domínio alfa e domínio do grupo de alta mobilidade (HMG), respectivamente. Em linhagens caracterizadas como *MAT1-1*, foram encontrados os genes *MAT1-1-1* e *MAT1-1-4*. Em linhagens caracterizadas como *MAT1-2*, foram encontrados os genes *MAT1-2-1* e *MAT1-2-5*. Com base nas análises das sequências obtidas de dez isolados de cada locus *MAT*, foi observado que ambos os idiomorfos possuem o gene da DNA Lyase (APN2) na região flangeadora 5', como esperado, demonstrando que ambos estão no mesmo locus, caracterizando assim, a espécie *P. citricarpa* como heterotática. Nesse trabalho ainda foi desenvolvido um protocolo de PCR Multiplex contendo *primers* específicos para amplificação do gene *MAT1-1-1* ou *MAT1-2-1*, que geram fragmentos de diferentes números de pares de base e possibilitam a classificação das linhagens como *MAT1-1* ou *MAT1-2* em uma única reação. Análises filogenéticas das sequências de nucleotídeos e proteínas dos loci *MAT* mostraram potencial desses genes na separação de diversos membros da classe Dotideomicetos.

**Palavras-chave:** *Phyllosticta citricarpa*, Mancha Preta dos Citros, reprodução sexual, genes *mating type*.



# IDENTIFICATION AND CHARACTERIZATION OF GENES INVOLVED IN THE SEXUAL REPRODUCTION OF THE ASCOMYCETE FUNGUS *Phyllosticta citricarpa*

## Abstract

The sexual reproduction in fungi is controlled by the *mating type* genes, located at *MAT* locus. In this study, we investigated the structure of this locus in the phytopathogenic fungus *Phyllosticta citricarpa*, causal agent of Citrus Black Spot. As with other heterothallic Ascomycetes, the idiomorphs present at the *MAT1* locus of this specie – *MAT1-1* and *MAT1-2* - encodes an alpha box protein and a high mobility group protein, respectively. In *MAT1-1* strains, we found *MAT1-1-1* and *MAT1-1-4* genes. In *MAT1-2* strains, we found *MAT1-2-1* and *MAT1-2-5* genes. Based on the analysis of the sequences obtained from ten isolates of each *MAT* locus, it was observed that both idiomorphs have the DNA Lyase gene (APN2) in the flanking region 5'. In this study was developed a Multiplex PCR protocol with specific *primers* to amplify *MAT1-1-1* or *MAT1-2-1* genes, generating fragments of different sizes and allows to classify the strains as *MAT1-1* or *MAT1-2* in a single reaction. Phylogenetic analysis of nucleotides and proteins sequences of *MAT* loci shown potential in the separation of species belonging to Dothideomycetes class.

**Keywords:** *Phyllosticta citricarpa*, Citrus Black Spot, sexual reproduction, *mating type* genes.

## 1. INTRODUÇÃO

A Mancha Preta dos Citros (MPC) é uma importante doença que afeta quase todas as variedades de laranja doce (*Citrus sinensis*) e limões (*C. limon*) (KOTZÉ, 1981) em quase todos os continentes, exceto na Europa, onde é considerada quarentenária A1 (EFSA, 2014). A MPC causa o aparecimento de lesões nas cascas dos frutos que o depreciam visualmente e diminuem sua aceitação no mercado *in natura*. Em estágios mais severos, causa a queda prematura dos frutos, acarretando em grandes prejuízos econômicos na produção (KOTZÉ, 1981).

O agente causal da MPC é o fungo *Phyllosticta citricarpa* e sua dispersão ocorre principalmente através dos esporos produzidos por reprodução sexuada – os ascósporos – que são carregados pelo vento e transmitem a doença a outras plantas de citros, que podem estar localizadas a grandes distâncias (FUNDECITROS, 2008). Embora a reprodução sexuada exerça um papel importante na epidemiologia dessa doença, seus mecanismos ainda não foram completamente elucidados. Para isso se faz necessário o estudo dos genes *mating type* presentes no locus *MAT*, que são responsáveis pelo controle da reprodução sexual em fungos (TSUI *et al.*, 2013).

Os idiomorfos *MAT* foram encontrados em dois genomas de *P. citricarpa*, utilizando-se as sequências dos idiomorfos e do gene APN2 (DNA Lyase) de *Diplodia pinea* (BIHON *et al.*, 2014) como referência. O gene APN2 já tinha sido anteriormente associado ao locus *MAT1* (DEBUCHY e TURGEON, 2006; BREWER *et al.*, 2011; TSUI *et al.*, 2013). A partir disso, foram selecionados *contigs* que possuíam altos níveis de similaridade com o gene APN2 de *Diplodia pinea*.

Neste trabalho, foi realizada a descrição da estrutura do locus *MAT* em *P. citricarpa* e demonstrado que, como acontece em outros Ascomicetos heterotáticos, esta espécie possui em seu locus *MAT* apenas um dos dois idiomorfos (*MAT1-1* que codifica para uma proteína de domínio alfa, ou *MAT1-2*, que codifica para uma proteína de domínio de alta mobilidade) (TURGEON e YODER, 2000) e necessita o encontro de indivíduos de *mating types* opostos para que ocorra o ciclo sexual. Assim, demonstramos pela primeira vez que a espécie *P. citricarpa* é heterotática em oposição a espécie proximamente relacionada e também encontrada em citros, *P. capitalensis*. Isso explica o fato de não ser possível a obtenção de ascósporos em condições de laboratório, quando trabalha-se com culturas puras do fungo *P. citricarpa* (BAAYERN

*et al.*, 2002, GLIENKE *et al.*, 2011). Essa pode ser a explicação para a dificuldade em reproduzir os sintomas de MPC em testes *in vitro*, uma vez que apenas conídios são inoculados. Desta forma, sugere-se a necessidade de utilização de ascósporos obtidos de linhagens *mating types* opostos para a realização do ciclo sexual *in vitro*.

Um protocolo de PCR Multiplex foi desenvolvido para facilitar o *screening* e estudos de distribuição e/ou incidência de diferentes *mating types* de *P. citricarpa* em determinada região, podendo assim prever o risco de ocorrência de reprodução sexual e disseminação da doença, além de facilitar estudos de cruzamento *in vitro*.

Análises filogenéticas baseadas em sequências de nucleotídeos e proteínas dos genes *MAT* mostraram-se eficientes na separação de espécies pertencentes à Classe Dotideomicetos.

## 2. MATERIAL E MÉTODOS

### 2.1 Linhagens utilizadas

Para o sequenciamento de DNA foram utilizadas vinte linhagens da espécie *Phyllosticta citricarpa* pertencentes à coleção do Laboratório de Genética de Microrganismos – LabGeM (Tabela 2).

### 2.2. Identificação do locus *MAT* a partir das sequências genômicas

Foram necessários genomas de duas linhagens diferentes, uma *MAT1-1* e outra *MAT1-2* para a identificação do locus *MAT* em *P. citricarpa*. As sequências genômicas utilizadas para identificação do gene *MAT1-2-1* foram obtidas no trabalho de Pacheco (2014).

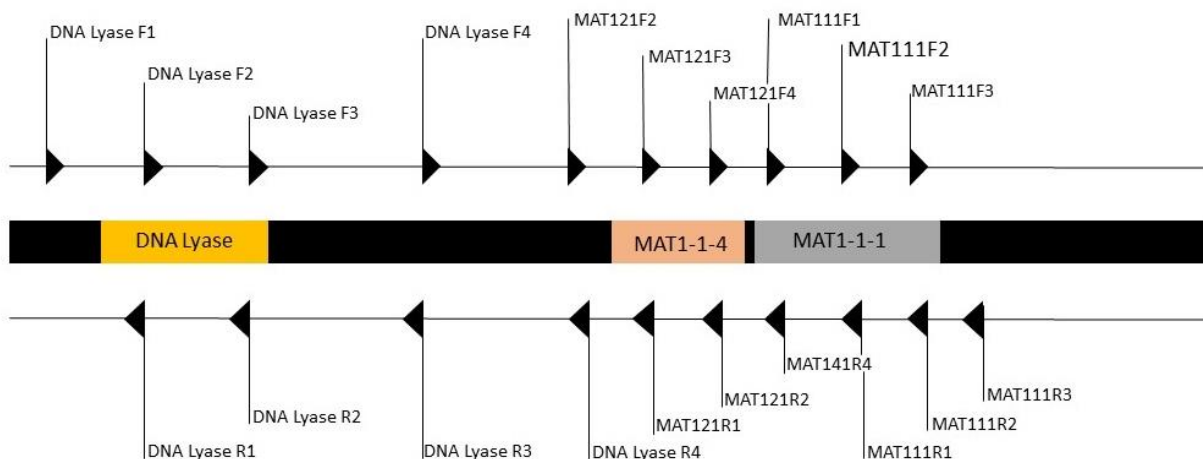
Para identificação do gene *MAT1-1-1*, foram utilizadas sequências genômicas de *P. citricarpa* disponíveis no GenBank (acesso: AOTE000000000.1).

Os idiomorfos *MAT* foram encontrados nesses genomas de *P. citricarpa*, utilizando-se as sequências dos idiomorfos e do gene APN2 (DNA Lyase) de *Diplodia pinea* (BIHON *et al.*, 2014) como referência. O gene APN2 já tinha sido anteriormente associado ao locus *MAT1* (DEBUCHY e TURGEON, 2006; BREWER *et al.*, 2011; TSUI *et al.*, 2013). A partir disso, foram selecionados *contigs* que possuíam altos níveis de similaridade com o gene APN2 de *Diplodia pinea*. Tais *contigs* foram então submetidos a predição de genes utilizando *software online* AUGUSTUS (<http://bioinf.uni-greifswald.de/webaugustus/>) e respectivas sequências de aminoácidos (Anexo 1). Estas sequências de aminoácidos foram então submetidas a análise através da plataforma *Basic Local Alignment Search Tool* (BLASTp) (ALTSCHUL *et al.*, 1997), disponível no site do NCBI – “National Center for Biotechnology Information”. Após a predição dos genes nestes *contigs*, tais sequências foram utilizadas para desenho de *primers* específicos para cada gene (Tabela 1).

### 2.2. Desenho de primers específicos para DNA Lyase, *MAT1-1-1*, *MAT1-2-1*

Pares de *primers* foram desenhados para cada um dos três genes através do Clone Manager Basic 9 – Sci-Ed Software ([http://www.sci-ed.com/pr\\_cmbas.htm](http://www.sci-ed.com/pr_cmbas.htm)) (Figura 2), sintetizados (Exxtend, Campinas, Brasil) e testados em diversos isolados de *P. citricarpa* pertencentes à coleção microbiológica do LabGeM.

### MAT1-1



### MAT1-2

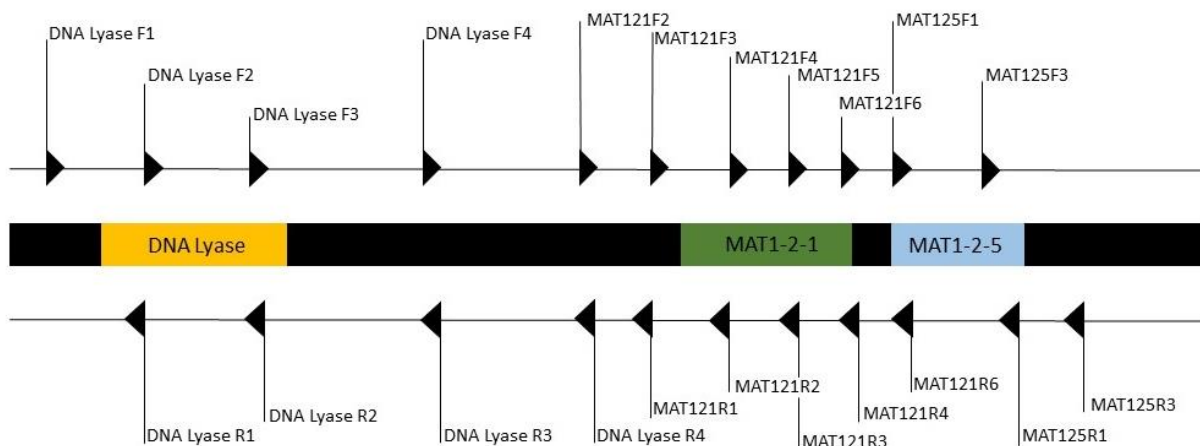


FIGURA 2: PRIMERS DESENHADOS PARA SEQUENCIAMENTO DA REGIÃO DO LOCUS *MAT1* DE *P. citricarpa*.  
FONTE: O Autor (2015)

## 2.3. Extração de DNA e amplificação por PCR

Para extração de DNA foram feitos de 15 a 20 inóculos em placas de Petri em papel celofane sobre meio Batata Dextrose Agar (BDA) Fluka Analytical® e a extração do DNA realizada através do UltraClean® Microbial DNA Kit – MoBio, utilizando protocolo do fabricante.

A reação de amplificação foi realizada contendo 0,5 U de TopTaq™ PCR Master Mix Quiagen, 0,2 µM de cada *primer*, 10 ng de DNA molde e água MilliQ para um

volume final de 12,5 µL. As sequências dos *primers* e ciclos utilizados na PCR estão descritos na Tabela 1.

As avaliações dos resultados foram realizadas em gel de agarose 1,5%, corado com GelRed®, visualizado sob luz UV (DNR Bio Imaging Systems Ltd.) e fotografado (GelCapture software).

#### **2.4. Seleção das linhagens de *P. citricarpa* para sequenciamento**

Foi realizada uma primeira seleção em 40 linhagens de *P. citricarpa* com os pares de *primers* MAT111F1 e MAT111R1 que amplificam exclusivamente linhagens *MAT1-1*, e MAT125F1 e MAT125R1 (Tabela 1) que amplificam exclusivamente linhagens *MAT1-2*. Dessas, foram selecionadas 20 linhagens: dez linhagens com o gene *MAT1-1-1* e dez linhagens com o gene *MAT1-2-1* para posterior sequenciamento.

#### **2.5. Sequenciamento**

Os produtos das PCR foram purificados com as enzimas Exo 01 e Fast AP (ThermoScientific kit). A reação de sequenciamento foi realizada com cada um dos pares de *primers* correspondentes, de acordo com protocolo ABI Big Dye® Terminator Cycle Sequencing Kit. A purificação do DNA foi feita por Sephadex™ G-50 DNA e as amostras foram então injetadas em sequenciador automático *ABI PRISM 3500 Genetic Analyzer*. Os gráficos gerados pelo sequenciador foram analisados e as sequências *forward* e *reverse* foram editadas para a obtenção de sequências consenso. Todas essas etapas foram realizadas através do software BioEdit.

#### **2.6. Alinhamento de nucleotídeos, proteínas e análise filogenética**

Sequências de nucleotídeos e de aminoácidos previstas através do software online AUGUSTUS (<http://bioinf.uni-greifswald.de/webaugustus/>) dos genes *MAT1-1-1* e *MAT1-2-1* de *P. citricarpa* foram alinhados com sequências correspondentes de outras espécies de ascomicetos disponíveis no GenBank utilizando MAFFT (Multiple Alignment using Fast Fourier Transform) para alinhamento de múltiplas sequências de aminoácidos e ácidos nucleicos (KATO e TOH, 2008).

Árvores filogenéticas foram geradas através do método Máxima Verossimilhança com 1000 replicações de *bootstrap* utilizando os softwares MEGA 6 (TAMURA *et al.*, 2013) e GARLI (BAZINET *et al.*, 2014).

## 2.7. PCR Multiplex para identificação dos idiomorfos *MAT1-1* e *MAT1-2*

Dentre os *primers* desenhados para identificação de linhagens com o gene *MAT1-1-1* e linhagens com o gene *MAT1-2-1*, foram selecionados os seguintes pares de *primers*: MAT111F3 (*forward*) 5' -GCA ATG TGG CAG CGC AAT CC-3' e MAT111R3 (*reverse*) 5' -GAT GAG TCC GAT GGT CCA GA-3' que amplificam exclusivamente o gene *MAT1-1-1* e MAT121F6 (*forward*) 5' - GAT CGT GGC AGG AGG CTT TG-3' e MAT121R6 (*reverse*) 5' - CTT ACC GAT CGC TGG TCG TT-3', que amplificam exclusivamente o gene *MAT1-2-1*. Cada par de primer foi selecionado por amplificar apenas um gene de um dos dois idiomorfos, por não formar dímeros e apresentar fragmentos de diferentes números de pares de base, possibilitando assim a caracterização como *MAT1-1* ou *MAT1-2* através da visualização em gel de agarose. A PCR foi realizada contendo 5,2 µL de água Milli-Q autoclavada, 0,5 U de TopTaq™ PCR Master Mix Quiagen e 0,2 µM de cada *primer*, com volume final de 12,5 µL e 10 ng de DNA. Os ciclos utilizados foram 94°C por 3 minutos, seguido por 30 ciclos com 30 segundos a 94°C, 30 segundos a 55°C, 1 minuto a 72°C, e uma extensão final por 10 minutos a 72°C

TABELA 1 – LISTA DE *PRIMERS* E SUAS RESPECTIVAS REGIÕES DE AMPLIFICAÇÃO, SEQUÊNCIAS E CICLOS UTILIZADOS PARA PCR

<b>Primers</b>	<b>Idiomorfo</b>	<b>Sequências</b>	<b>Ciclo</b>
<b>DNA Lyase F1</b> <b>DNA Lyase R1</b>	<i>MAT1-1 e MAT1-2</i>	5' ATGCCGCTCCGGATAACCTC 3' 3' GAGGCCCTCCTGAATGTTGG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 57°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>DNA Lyase F2</b> <b>DNA Lyase R2</b>	<i>MAT1-1 e MAT1-2</i>	5' TGGCTCCAGGATCGACTATG 3' 3' GACGAGTTGGAGGCTGATTC 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 54°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>DNA Lyase F3</b> <b>DNA Lyase R3</b>	<i>MAT1-1 e MAT1-2</i>	5' CTACCGCTCTACCGCAATCG 3' 3' CGTGTGCCACCAGTTTCTTC 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>DNA Lyase F4</b> <b>DNA Lyase R4</b>	<i>MAT1-1 e MAT1-2</i>	5' GCGCATGAATCAAGACAACG 3' 3' GCTCCTCCTCAGAGAACAAG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT121 F2</b> <b>MAT121 R1</b>	<i>MAT1-1 e MAT1-2</i>	5' CCTGCACAACCTGCATCTTC 3' 3' TCGCTTTGGCACAAGGTACG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 57°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT121 F3</b> <b>MAT121 R2</b>	<i>MAT1-1 e MAT1-2</i>	5' TCTTCTGCTATGCCGAGACG 3' 3' GGATTCACAGGCTCACAAGG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 56°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT121 F4</b> <b>MAT121 R3</b>	<i>MAT1-2</i>	5' CGCGGTTCGATGTGGGAATTG 3' 3' CCGGCTATTTCTGTGAGAC 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 54°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT121 F5</b> <b>MAT121 R4</b>	<i>MAT1-2</i>	5' GGAGGACAGGGTGACAAGAG 3' 3' GTCCGCGGTTCGACAAGAATG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT121 F6</b> <b>MAT121 R6</b>	<i>MAT1-2</i>	5' GATCGTGGCAGGAGGCTTTG 3' 3' AACGACCAGCGATCGGTAAG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 56°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT125 F1</b> <b>MAT125 R1</b>	<i>MAT1-2</i>	5' TCCACCGATTCTCTTCTCAC 3' 3' CGACGAGCTTGACGAGATTG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 52°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT125 F3</b> <b>MAT125 R3</b>	<i>MAT1-2</i>	5' CGCGAAGCTGCAATCTCGTC 3' 3' CAGACTGAGAGGGCCTTCTG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'



<b>MAT121F4</b> <b>MAT141 R4</b>	<i>MAT1-1</i>	5' CGCGGTCGATGTGGGAATTG 3' 3' TGAGCCGAAGAACAAGTTGG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT111 F1</b> <b>MAT111 R1</b>	<i>MAT1-1</i>	5' ATGGCTGTCGCTGCAAACCTTC 3' 3' CCCAAGGTGACAAAGTATGC 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT111 F2</b> <b>MAT111 R2</b>	<i>MAT1-1</i>	5' TCCTCGCTCGAGATGCATAC 3' 3' AGAGCCTGGTCATTGGATTG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 54°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'
<b>MAT111 F3</b> <b>MAT111 R4</b>	<i>MAT1-1</i>	5' GCAATGTGGCAGCGCAATCC 3' 3' ACTTCAAATCCGACCGACTG 5'	94°C por 1', seguido por 30 ciclos de 94°C por 30", 55°C por 40" e 72°C por 40" e uma extensão final de 72°C por 7'

### 3. RESULTADOS E DISCUSSÃO

#### 3.1. Identificação do locus *MAT*

Para a identificação dos locus *MAT* foram utilizados os genomas de duas linhagens de *P. citricarpa*. Para tanto, utilizou-se a sequência do gene APN2 de *Diplodia pinea* como referência. Uma sequência com homologia a este gene foi encontrada no *contig* 4071 (acesso: AOTE01003969) de aproximadamente 9 kb, do genoma de *P. citricarpa* CGMCC3.14348 (acesso: AOTE00000000.1) e no *contig* 201 da linhagem LGMF05 de aproximadamente 175 kb. Tais *contigs* foram analisados através da plataforma *Basic Local Alignment Search Tool* (BLASTx) (ALTSCHUL *et al.*, 1997), disponível no site do NCBI. Foi observado um fragmento de 317 pb similar ao gene *MAT1-1-1* de *Diplodia pinea* no genoma de *P. citricarpa* CGMCC3.14348. Não foram identificados genes com similaridade ao gene *MAT1-2-1* no genoma desse isolado. Por outro lado, foi identificado no *contig* 201 da linhagem LGMF05 uma região de similaridade com o gene *MAT1-2-1* que codifica para a proteína que contém HMG box. Não foi encontrada nenhuma região de homologia com o gene *MAT1-1-1* no genoma desta linhagem. Estas sequências foram utilizadas para o desenho de *primers* para a caracterização do locus *MAT1* de *P. citricarpa*.

#### 3.2. Sequenciamento de linhagens com os idiomorfos *MAT1-1* e *MAT1-2*

Para a reação de sequenciamento foram selecionadas dez linhagens *MAT1-1* e dez linhagens *MAT1-2* dentro da coleção microbiológica do LabGeM, as quais estão indicadas com asterisco na Tabela 2. As linhagens previamente caracterizadas como *MAT1-1* ou *MAT1-2* através da ferramenta Blastx tiveram o locus *MAT* confirmado por sequenciamento. Não foram observadas diferenças nucleotídicas nas sequências de linhagens pertencentes a um mesmo idiomorfo *MAT*. Resultados semelhantes já foram descritos na literatura para outros fungos, onde não foram encontradas diferenças intra-específicas significativas (BENNET *et al.*, 2003; TSUI *et al.*, 2013).

TABELA 2: LINHAGENS DE *Phyllosticta citricarpa* UTILIZADAS NESSE ESTUDO, MATING TYPE, CÓDIGO DA LINHAGEM, ORIGEM, TECIDO VEGETAL DE ISOLAMENTO E HOSPEDEIRO

Mating Type	Linhagem	Origem	Substrato	Hospedeiro
MAT1.1	LGMF 15	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 16	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 24	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 32	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 33	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 34	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 41	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 49	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 50	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 53	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 56	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 59	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 63	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 66	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 78*	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 83*	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 89*	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 90*	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 92	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 98	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 103	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 105	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 108	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 109	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 111	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 113	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 115	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>

	LGMF 134*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 145*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 146	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 159*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 162*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 170*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 179	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 193	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 216	Ribeirão do Scheffer (Paraná, Brasil)	Folha	Laranja; <i>Citrus sinensis</i>
	LGMF 309*	Quarteirão dos Órfãos (Paraná, Brasil)	Folha	Tangerina; <i>Citrus reticulata</i>
	LGMF 312	Quarteirão dos Órfãos (Paraná, Brasil)	Folha	Tangerina; <i>Citrus reticulata</i>
<b>MAT 1.2</b>	LGMF 04	Conchal (São Paulo, Brasil)	Folha	Lima Tahiti; <i>Citrus latifolia</i>
	LGMF 05*	Conchal (São Paulo, Brasil)	Fruto	Lima Tahiti; <i>Citrus latifolia</i>
	LGMF 06*	Conchal (São Paulo, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 20	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 22	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 25	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 28	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 30*	Cerro Azul (Paraná, Brasil)	Fruto	Laranja; <i>Citrus sinensis</i>
	LGMF 44	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 45	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 52	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 54*	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 55	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 61*	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 64	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 65	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 69	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 75	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
	LGMF 81	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>

LGMF 86	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 88	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 91	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 96	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 97	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 102*	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 104	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 106	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 114*	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 122*	Ribeirão do Scheffer (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 127*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 143	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 144	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 153	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 161	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 169	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 171*	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 180	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 183	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 184	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 187	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 194	Cerro Azul (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 202	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 203	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 204	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 205	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 206	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 209	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>

LGMF 210	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 211	Dr. Ulysses (Paraná, Brasil)	Fruto	Tangerina; <i>Citrus reticulata</i>
LGMF 241	Ribeirão do Scheffer (Paraná, Brasil)	Folha	Limão cravo; <i>Citrus limonia</i>
LGMF 246	Ribeirão do Scheffer (Paraná, Brasil)	Folha	Limão cravo; <i>Citrus limonia</i>
LGMF 280	Ribeirão do Scheffer (Paraná, Brasil)	Folha	Tangerina; <i>Citrus reticulata</i>
LGMF 283	Quarteirão dos Órfãos (Paraná, Brasil)	Folha	Tangerina; <i>Citrus reticulata</i>

\* Linhagens utilizadas para o sequenciamento.

### 3.3. Estrutura do locus *MAT1* e suas regiões flangeadoras

O comprimento dos genes *mating type* foram previstos através do programa FGENESH+ *gene prediction* (SOLOVYEV, 2007). O gene *MAT1-1-1* de *P. citricarpa* possui 1500 pb (499 aa) e está no mesmo locus do gene *MAT1-1-4* (873 pb e 290 aa), semelhante ao observado em *Diplodia pinea* (BIHON *et al.*, 2014). Segundo Groenewald (2007), a proteína codificada pelo gene *MAT1-1-4* é da família das metalotioneínas, as quais são pequenas proteínas ligadoras de metal encontradas em todos os eucariotos e em vários procariotos (BLINDAUER *et al.*, 2001; COYLE *et al.*, 2002). Essa proteína já foi descrita como necessária para o normal funcionamento dos apressórios e à patogenicidade em *Magnaporthe grisea* (TUCKER *et al.*, 2004). Além disso, foi sugerido seu envolvimento na via de especialização da célula e morfogênese sexual em *Pyrenopeziza brassicae* (SINGH e ASHBY, 1998). Porém, a importância dessa proteína no desenvolvimento sexual de fungos ainda não está completamente elucidada.

O gene *MAT1-2-1* de *P. citricarpa* possui 1218 pb e codifica para uma proteína de 405 aminoácidos. Ainda no mesmo locus foi demonstrada a presença do gene *MAT1-2-5* com um tamanho de 1101 pb (323 aa). Dados semelhantes foram encontrados em *Diplodia pinea* (BIHON *et al.*, 2014) e em *Pseudogymnoascus destructans*, fungo causador da Síndrome do Nariz Branco em morcegos (PALMER *et al.*, 2014). A busca realizada utilizando a ferramenta BLASTx com o gene *MAT1-2-5* não mostrou homologia significativa com proteínas previamente descritas. Esse gene recebeu tal denominação porque foi o quinto a ser descoberto e associado ao idiomorfo *MAT1-2* (BIHON *et al.*, 2014).

O gene da enzima DNA Lyase (APN2) foi encontrado em ambos os genomas de *P. citricarpa*. Esse gene se encontra na região flangeadora 5' muito próximo ao locus *MAT*, como previamente descrito para a maioria dos Doideomicetos heterotáticos (DEBUCHY *et al.*, 2010). O mesmo não é observado em *Diplodia pinea*, na qual existem 5 genes entre o locus *MAT* e o gene da DNA Lyase, resultando em aproximadamente 13 kb de distância entre ambos (BIHON *et al.*, 2014). Em *Grosmannia clavigera*, um importante patógeno de coníferas, o gene APN2 está 15 kb distante do locus *MAT* (TSUI *et al.*, 2013).

A presença exclusiva de um idiomorfo *MAT* (*MAT1-1* ou *MAT1-2*) no mesmo locus nos genomas das linhagens de *P. citricarpa* estudadas, evidenciam essa espécie como heterotática. Sendo assim, para que ocorra o ciclo sexual nesse fungo, se faz necessária a presença de indivíduos de *mating types* opostos.

A organização do gene da DNA Lyase e dos genes do locus *MAT* pode ser observada na Figura 3.

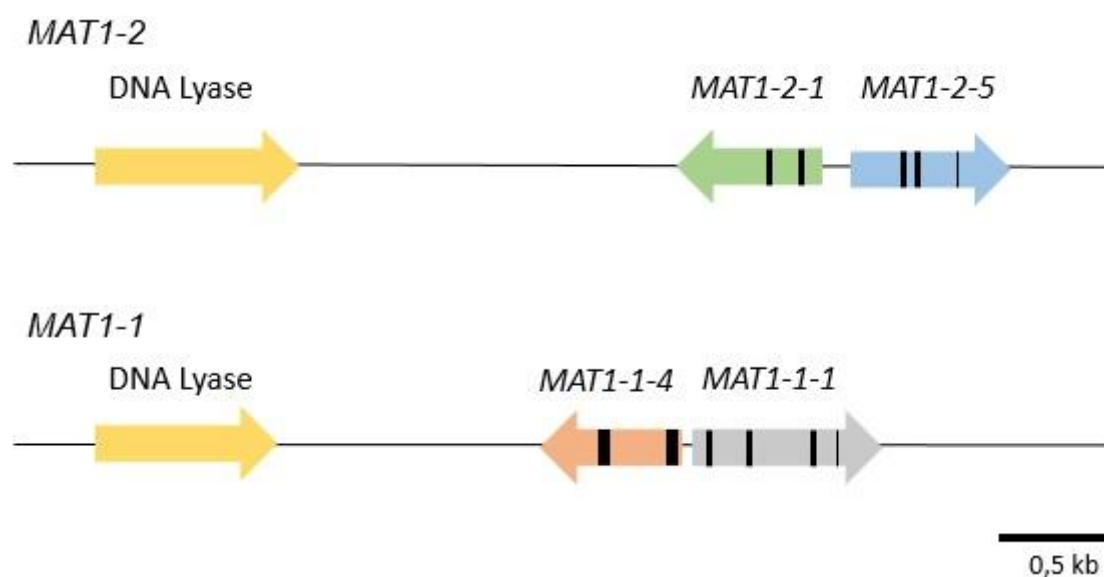


FIGURA 3: ORGANIZAÇÃO DO LOCUS *MAT1* NA ESPÉCIE HETEROTÁTICA *P. citricarpa*.

FONTE: O Autor (2015)

Nota: A barra de escala indica o tamanho dos idiomorfos *MAT*, bem como de seu gene flangeador *DNA Lyase* e a distância entre eles. Setas indicam o sentido de transcrição. Traços pretos indicam íntrons.

### 3.4. PCR Multiplex para identificação dos idiomorfos *MAT1-1* e *MAT1-2*

Um protocolo de PCR Multiplex utilizando os *primers* MAT111F3, MAT111R3, MAT121F6 e MAT121R6 (Tabela 1) para identificação rápida e eficiente de cada idiomorfo *MAT* foi desenvolvido neste trabalho. Como pode-se observar na figura 4, linhagens que possuem o gene *MAT1-1-1* amplificam um fragmento de 606 pb, enquanto que linhagens que possuem o gene *MAT1-2-1* amplificam um fragmento de 692 pb. Das 91 linhagens testadas, 100% apresentaram bandas visíveis de 606 ou 692 pb. A PCR Multiplex já foi utilizada anteriormente com sucesso na identificação de genes *mating type* para os fungos heterotáticos *Tuber melanosporum* (RUBINI *et al.*, 2010), *Tapesia yallundae* e *Tapesia acuformis* (DYER *et al.*, 2001). No trabalho de Dyer



e colaboradores (2001) foi utilizado um *primer forward* comum para os idiomorfos *MAT1-1* e *MAT1-2* em *T. yallundae* e dois *primers* específicos para as sequências dos idiomorfos *MAT1-1* ou *MAT1-2*. Porém, em *T. acuformis* houve fraca ou nenhuma amplificação utilizando-se os mesmos pares de *primers*, corroborando com estudos que afirmam a grande dissimilaridade nas sequências dos genes *mating type* entre espécies (LI *et al.*, 2013), inclusive entre espécies do mesmo gênero. Além disso, PCR Multiplex para os genes *mating type* também já foi utilizada na identificação do fungo *Villosiclava virens* como homotático, por possuir tanto o gene *MAT1-1-1* quanto *MAT1-2-1* em seu genoma, apresentando ambas as bandas de 256 e 222 pb em 90% dos isolados testados (FU *et al.*, 2014).

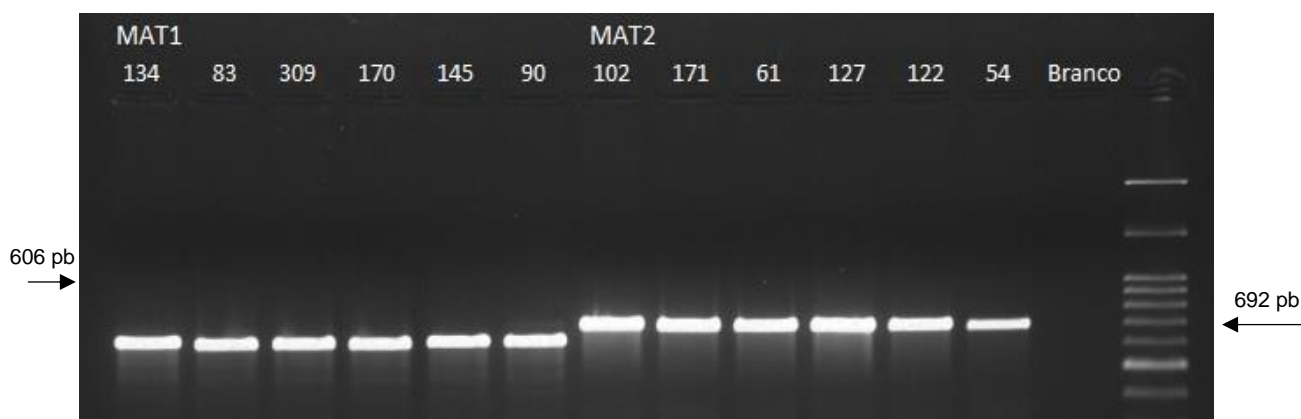


FIGURA 4: PRODUTOS DE AMPLIFICAÇÃO DE PCR MULTIPLEX PARA LINHAGENS DE *P. citricarpa* *MAT1-1* (606 PB) E *MAT1-2* (692 PB).

FONTE: O Autor (2015)

O protocolo de PCR Multiplex *mating type* específico desenvolvido nesse trabalho foi capaz de identificar as linhagens de *P. citricarpa* testadas como *MAT1-1* ou *MAT1-2*, gerando amplificações de fragmentos com diferentes números de pares de base em uma mesma reação e facilmente identificáveis em gel de agarose. Foram testadas 91 linhagens de diversos locais de origem (Tabela 2), sendo que 58,24% foram caracterizadas como *MAT1-2* e 41,76% como *MAT1-1*. Também foram analisadas as proporções das populações de cada um dos pomares separadamente e análises de qui-quadrado indicam que a relação de *MAT1-1* e *MAT1-2* não é significativamente diferente de 1:1 ( $P < 0,05$ ), como mostra a figura 5.

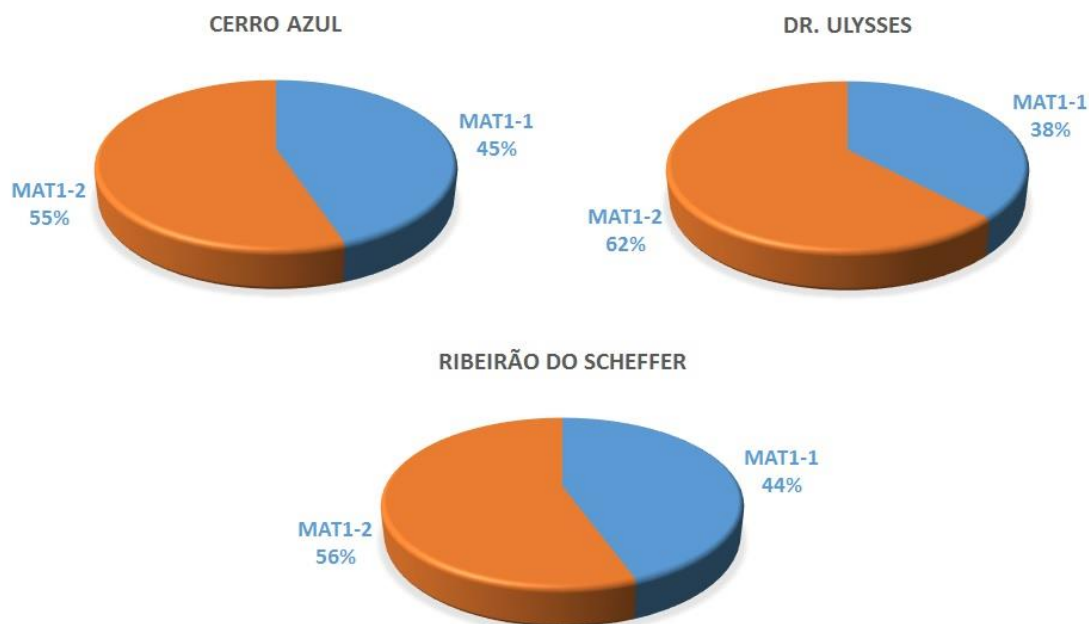


FIGURA 5: GRÁFICOS MOSTRANDO AS PROPORÇÕES DOS IDIOMORFOS *MAT1-1* E *MAT1-2* NOS POMARES DE CERRO AZUL, DR. ULYSSES E RIBEIRÃO DO SCHEFFER.  
 FONTE: O Autor (2015)

A PCR Multiplex para os genes *mating type* pode ser utilizada em estudos de distribuição e/ou incidência de diferentes *mating types* em determinada região prevendo o risco de ocorrência de reprodução sexual, além de facilitar estudos de cruzamento *in vitro*. Outra aplicação importante é a seleção de linhagens de diferentes *mating types* para a produção de ascósporos em experimentos de indução de sintomas de MPC *in vitro*.

### 3.5. Árvores filogenéticas baseadas em sequências de genes *mating type*

Para as análises filogenéticas foram utilizadas sequências de nucleotídeos e aminoácidos dos genes *mating type* de fungos fitopatogênicos pertencentes a diversas famílias da classe Dotideomicetos. As árvores filogenéticas geradas utilizando sequências de aminoácidos (Figura 6) e nucleotídeos (Figura 7) do gene *MAT1-1-1* apresentaram a mesma topologia e ambas estão em concordância com a filogenia de Dotideomicetos descrita por Schoch e colaboradores (2009). Em todas as análises pode-se observar um clado formado pelas linhagens *Pseudocercospora musae*, *Pseudosercospora fijiensis*, *Dothistroma septosporum*, *Cercospora beticola* e

*Mycosphaerella graminicola*, todos pertencentes a família *Mycosphaerellaceae*. O segundo clado é formado pelas linhagens *Diplodia pinea* e *Phyllosticta citricarpa* pertencentes às famílias *Botryosphaeriaceae* e *Phyllostictaceae*, respectivamente, e as últimas duas linhagens, *Cochilobolus heterostrophus* e *Alternaria alternata*, à família *Pleosporaceae*. Uma pequena divergência foi observada na árvore filogenética gerada a partir de sequências nucleotídicas, onde a linhagem de *Cercospora beticola* está localizada mais externa ao clado da família *Mycosphaerellaceae*.

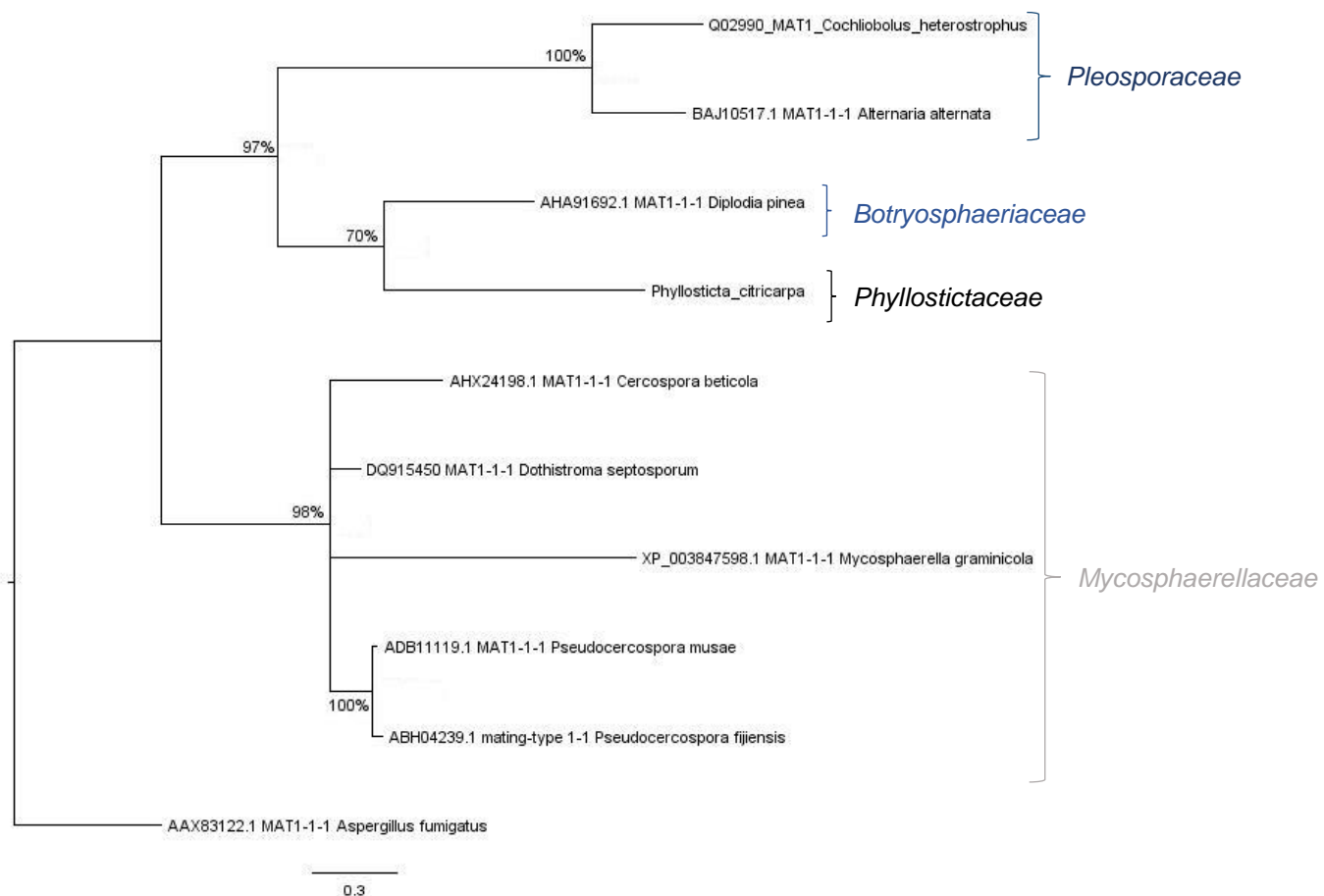


FIGURA 6: ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE AMINOÁCIDOS DO GENE *MAT1-1-1*.

Nota: Porcentagem de bootstrap com 1000 reamostragens através de análise de Máxima Verossimilhança estão indicados a esquerda dos nós. A árvore foi enraizada com *Aspergillus fumigatus* (AAX83122).

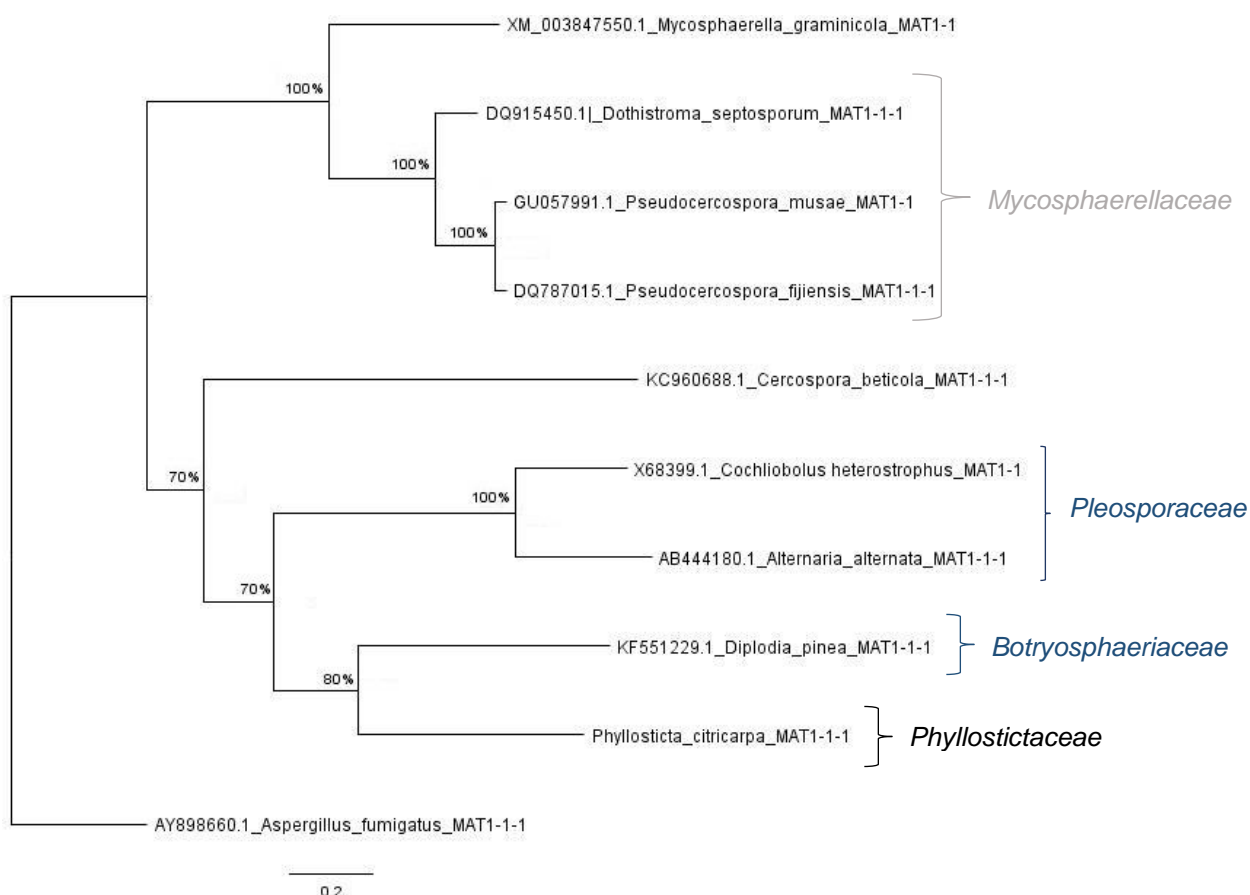


FIGURA 7: ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE NUCLEOTÍDEOS DO GENE *MAT1-1-1*.

Nota: Porcentagem de bootstrap com 1000 reamostragens através de análise de Máxima Verossimilhança estão indicados a esquerda dos nós. A árvore foi enraizada com *Aspergillus fumigatus* (AY898660).

As árvores filogenéticas geradas a partir de sequências de aminoácidos (Figura 8) e nucleotídeos (Figura 9) do gene *MAT1-2-1* mostraram a mesma topologia e ambas estão em concordância com a filogenia de Dothideomicetos descrita por Schoch e colaboradores (2009). *Cercospora beticola* está novamente alocada como uma espécie mais externa ao clado da família *Mycosphaerellaceae* na árvore filogenética gerada a partir de sequências de nucleotídeos, semelhante ao ocorrido com essa espécie na análise com o gene *MAT1-1-1*.

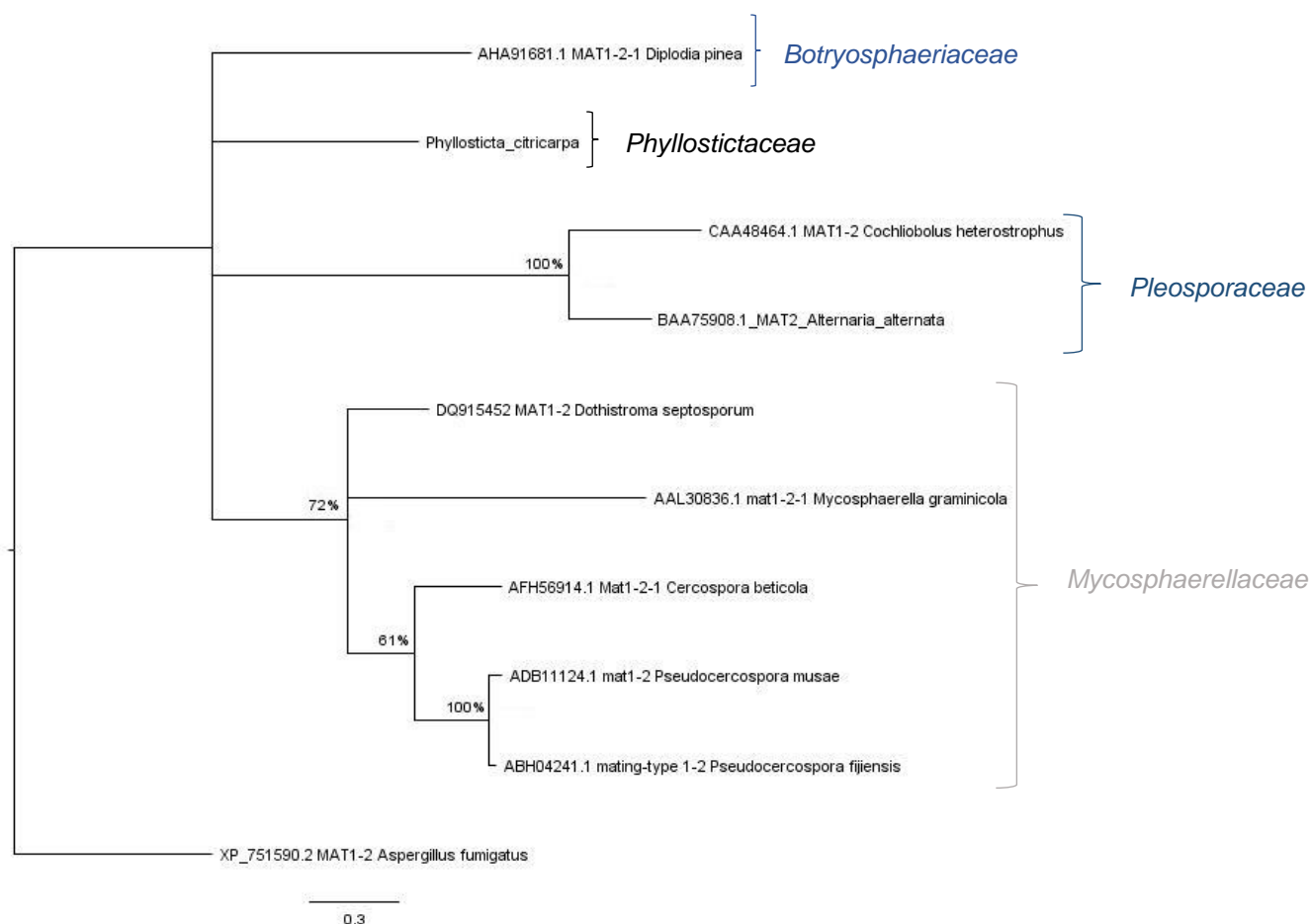


FIGURA 8: ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE AMINOÁCIDOS DO GENE *MAT1-2-1*.

Nota: Porcentagem de bootstrap com 1000 reamostragens através de análise de Máxima Verossimilhança estão indicados a esquerda dos nós. A árvore foi enraizada com *Aspergillus fumigatus* (XP\_751590).

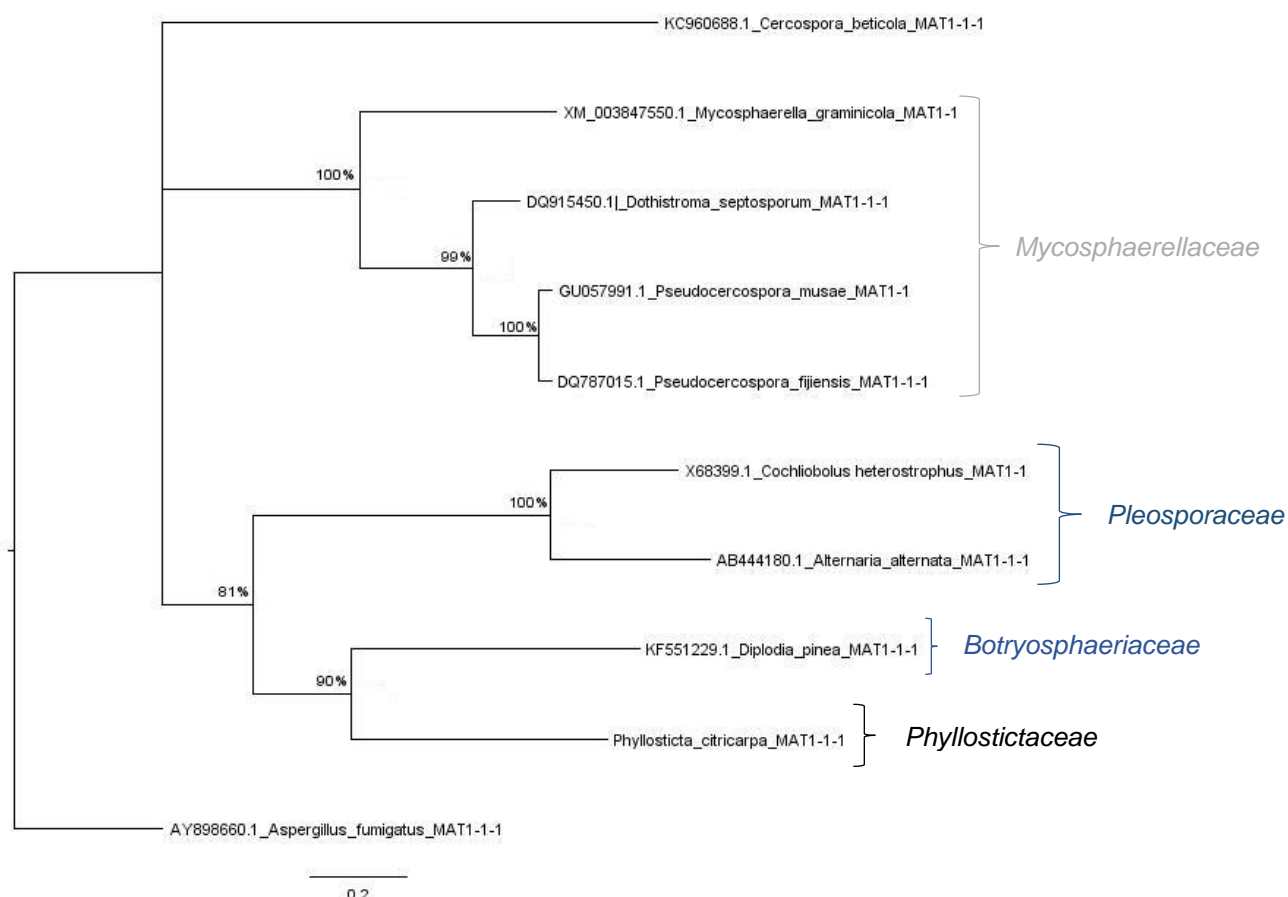


FIGURA 9: ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE NUCLEOTÍDEOS DO GENE *MAT1-2-1*.

Nota: Porcentagem de bootstrap com 1000 reamostragens através de análise de Máxima Verossimilhança estão indicados a esquerda dos nós. A árvore foi enraizada com *Aspergillus fumigatus* (AY898660).

Análises filogenéticas das sequências de nucleotídeos do gene DNA Lyase e de aminoácidos da proteína correspondente não mostraram diferenças entre linhagens caracterizadas como *MAT1-1* e *MAT1-2*. Sendo assim, demonstramos que as cópias do gene APN2 nas duas linhagens de *P. citricarpa* (*MAT1-1* e *MAT1-2*) são idênticas e são ortólogas em relação aos demais ascomicetos analisados aqui. Ainda, a análise filogenética deste gene mostrou-se eficaz para a classificação em nível de gênero, porém não apresenta correlação filogenética entre os idiomorfos *MAT* (Figuras 10 e 11). O gene da DNA Lyase também mostrou topologia de árvore similar à observada por Schoch e colaboradores (2009) na árvore filogenética da Classe Dotideomicetos, porém, na análise filogenética baseada em sequências de aminoácidos a linhagem *Mycosphaerella graminicola* mostrou-se mais correlacionada com espécies

pertencentes as famílias *Phyllostictaceae* e *Botryosphaeriaceae*, não agrupando no mesmo clado da família *Mycosphaerellaceae*.

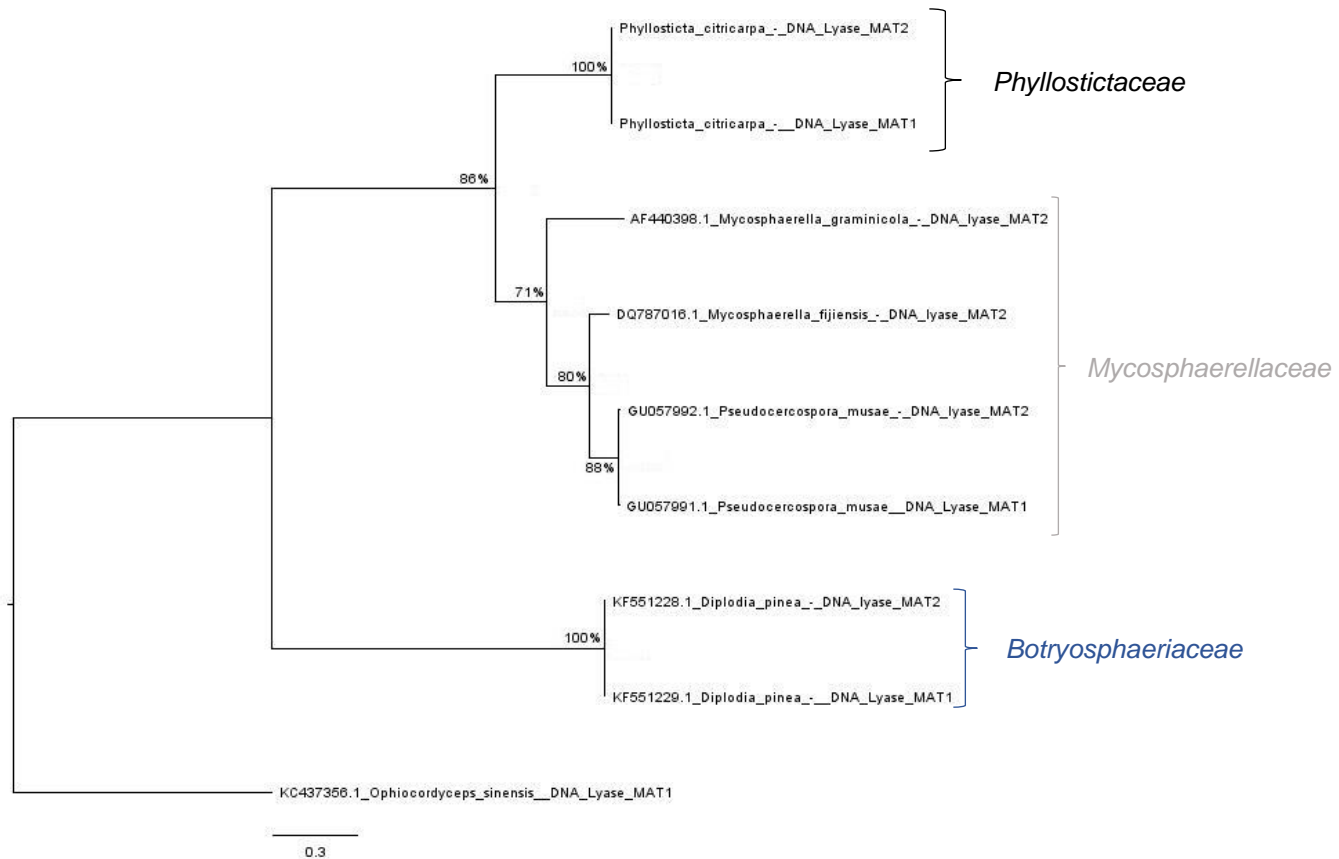


FIGURA 10: ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE NUCLEOTÍDEOS DA ENZIMA DNA Lyase.

Nota: Porcentagem de bootstrap com 1000 reamostragens a partir de análise de Máxima Verossimilhança estão indicados a esquerda dos nós. A árvore foi enraizada com *Ophiocordyceps sinensis* (KC437356).



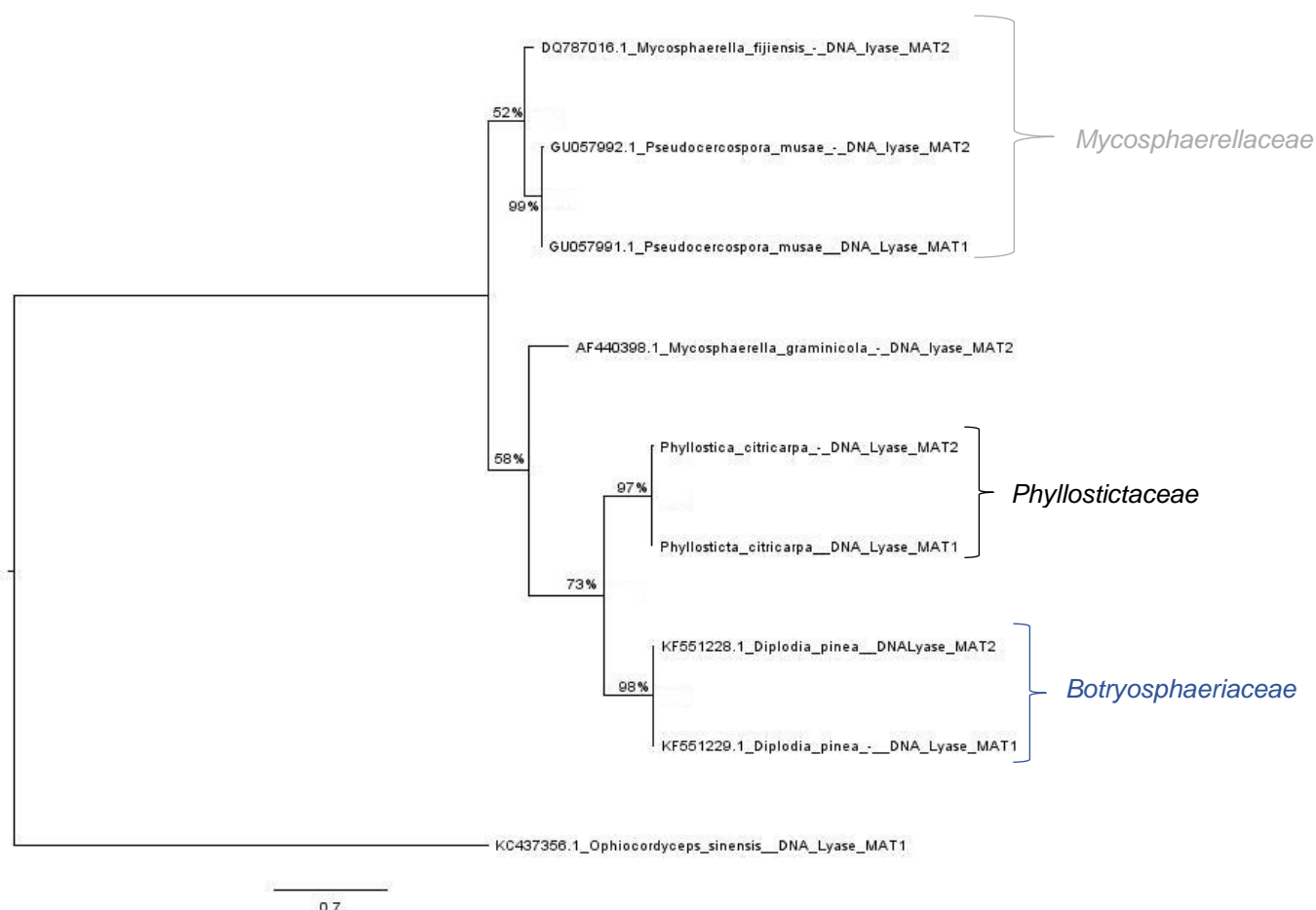


FIGURA 11: ANÁLISE FILOGENÉTICA DE SEQUÊNCIAS DE AMINOÁCIDOS DA ENZIMA DNA Lyase.

Nota: Porcentagem de bootstrap com 1000 reamostragens a partir de análise de Máxima Verossimilhança estão indicados a esquerda dos nós. A árvore foi enraizada com *Ophiocordyceps sinensis* (KC437356).

Como já dito anteriormente, o termo idiomorfo é utilizado para designar sequências de genes que ocupam a mesma posição no cromossomo, porém são estruturalmente diferentes, codificam para diferentes fatores de transcrição e normalmente não se recombina, fazendo com que sejam caracteres filogeneticamente independentes (KOHMOTO e YODER, 1997). Vários autores defendem que os genes *mating type* tem grande potencial como marcadores moleculares para análises filogenéticas (YUN *et al.*, 2000; PÖGGELER, 2001; DU *et al.*, 2005), inclusive comparando-os a sequências *Internal Transcribed Spacer* (ITS) (DU *et al.*, 2005), que são amplamente utilizadas em estudos filogenéticos de plantas e fungos devido a fatores como universalidade, uniformidade intragenômica, variabilidade intergenômica, baixa restrição funcional, entre outros (ÁLVAREZ e WENDEL, 2003). Há a hipótese de que os genes *mating type* sofram pressão de seleção pela necessidade de ser mantida

a função das proteínas ligadas ao desenvolvimento sexual (DU *et al.*, 2005). Entretanto, certas regiões são conservadas enquanto outras são altamente variáveis e, além disso, a taxa de variação inter-específica é alta, enquanto que a taxa de variação intra-específica é baixa (TURGEON, 1998) podendo assim ser úteis na resolução de relações entre organismos mais intimamente relacionados (KOHMOTO e YODER, 1997).

A topologia da árvore filogenética baseada em sequências de aminoácidos codificados pelos genes *MAT1-1-1* e *MAT1-2-1* está de acordo com a filogenia dos Dotideomicetos descrita por Schoch e colaboradores (2009). Porém, há divergência no posicionamento da linhagem de *Cercospora beticola* nas árvores geradas a partir de sequências de nucleotídeos e proteínas. Na árvore baseada em nucleotídeos, a linhagem *C. beticola* está alocada externamente ao clado da família *Mycosphaerellaceae*. Isso pode ser explicado pelo fato da região variável do gene *MAT1-2-1* dessa espécie ser divergente quando comparada às sequências nucleotídicas dos outros membros da família *Mycosphaerellaceae*, porém, sua região conservada codifica a sequência de proteínas responsável por manter as características da família *Mycosphaerellaceae* em relação ao seu ciclo sexual.

Diversos estudos vêm corroborando com a afirmação de que os genes *mating type* podem ser úteis na delimitação a nível de espécies devido a presença de uma região variável, a qual sofre processo de rápida evolução nos genomas. Du e colaboradores (2005) utilizaram o gene *MAT1-2-1* na resolução de espécies do complexo *Collectotrichum*, e concluíram que os genes *mating type* apresentaram melhor resolução que sequências da região ITS. O'Donnell e colaboradores (2004) utilizaram os genes *mating type* no reconhecimento de nove novas espécies filogeneticamente distintas de fungos produtores de toxinas tricotecenos no clado de *Fusarium graminearum*. Os autores ainda afirmam que os genes *mating type* se tornam ainda mais informativos quando incluídas também nas análises as regiões intergênicas. Membros da família *Clavicipitaceae*, pertencente ao filo Ascomycota, foram submetidos à análise filogenética utilizando os genes 18s rRNA, *MAT1-1-1* e *MAT1-2-1*. As filogenias dos genes *mating type* mostraram-se melhores nas resoluções dentro dessa família que sequências do gene 18s rRNA (YOKOYAMA *et al.*, 2006).

Com base nos dados observados, foi elucidada a estrutura do locus *MAT* de *P. citricarpa* e demonstrado que esse fungo fitopatogênico é heterotático, apresentando em seu genoma os idiomorfos *MAT1-1* ou *MAT1-2*. Foi desenvolvido um protocolo de PCR Multiplex para facilitar a identificação dos idiomorfos *MAT1-1* e *MAT1-2* de linhagens de *P. citricarpa* em uma única reação, gerando amplificações de fragmentos com diferentes números de pares de base. Análises filogenéticas realizadas a partir de sequências de nucleotídeos e aminoácidos dos genes *mating type* e DNA Lyase mostraram-se eficientes na separação das espécies selecionadas, sendo que análises utilizando sequências de aminoácidos e nucleotídeos apresentaram topologias semelhantes. As árvores utilizando sequências de aminoácidos mostraram um suporte maior para a separação em nível de família, em contrapartida, as árvores baseadas em sequências nucleotídicas apresentaram maior suporte de bootstrap nos ramos internos – separação em nível de gênero/espécie.

## CONCLUSÕES

- O fungo *Phyllosticta citricarpa* é heterotático devido a presença dos idiomorfos *MAT1-1* ou *MAT1-2* no locus *MAT1*.
- Foi descrita a estrutura do locus *MAT1* de *Phyllosticta citricarpa* em linhagens que contém os idiomorfos *MAT1-1* e em linhagens que contém *MAT1-2*. Em *P. citricarpa* idiomorfo *MAT1-1*, estão presentes os genes *MAT1-1-1* e *MAT1-1-4* próximos à região 3' do gene da enzima DNA Lyase. Em *P. citricarpa* idiomorfo *MAT1-2*, estão presentes os genes *MAT1-2-1* e *MAT1-2-5*, sendo que nessas linhagens, os genes também estão localizados próximos à região 3' do gene da enzima DNA Lyase.
- O protocolo de PCR Multiplex desenvolvido nesse trabalho mostrou-se eficiente na determinação e diferenciação dos idiomorfos *MAT1-1* ou *MAT1-2* de *P. citricarpa*.
- As linhagens *MAT1-1* ou *MAT1-2* encontram-se em proporções semelhantes na coleção de cultura do LabGeM.
- As análises filogenéticas realizadas com sequências de nucleotídeos e aminoácidos dos genes *mating type* apresentaram poder de resolução na classificação entre espécies pertencentes a classe Dotideomicetos.

#### 4. REFERÊNCIAS BIBLIOGRÁFICAS

ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A A; et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic acids research**, v. 25, n. 17, p. 3389–3402, 1997. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=146917&tool=pmcentrez&rendertype=abstract>>.

ÁLVAREZ, I.; WENDEL, J. F. Ribosomal ITS sequences and plant phylogenetic inference. **Molecular Phylogenetics and Evolution**, v. 29, p. 417-434. 2003.

BAAYEN, R. P.; BONANTS, P. J. M.; VERKLEY, G.; CARROLL, G. C.; VAN DER Aa, H. A.; WEERDT, M de; VAN BROUWERSHAVEN, I. R.; SCHUTTE, G. C.; MACCHERONI JR, W.; GLIENKE DE BLANCO, C.; AZEVEDO, J. L. Nonpathogenic Isolates of the Citrus Black Spot Fungus, *Guignardia citricarpa*, Identified as a Cosmopolitan Endophyte of Woody Plants, *G. mangiferae* (*Phyllosticta capitalensis*). **Ecology and Population Biology**. V. 92, n. 5. 2002.

BAZINET; ZWICKI; CUMMINGS. The phylogenetic analyses performed by the GARLI web service hosted at [molecularrevolution.org](http://molecularrevolution.org) used GARLI 2.1 (Genetic Algorithm for Rapid Likelihood Inference; Zwickl) 2006.

BIHON, W.; WINGFIELD, M. J.; SLIPPERS, B.; DUONG, T. A.; WINGFIELD, B. D. MAT gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. **Fungal Genetics and Biology**, v. 62, p. 55-61. 2014.

BLINDAUER, C.A.; HARRISON, M.D.; PARKINSON, J.A.; ROBINSON, A.K.; CAVET, J.S.; ROBINSON, N.J.; SADLER, P.J. A metallothionein containing zinc finger within a four-metal cluster protects bacterium from zinc toxicity. **Proc. Natl. Acad. Sci. USA** 98, 9593– 9598. 2001.

BREWER, M. T.; CADLE-DAVIDSON, L.; CORTESI, P.; SPANU, P. D.; MILGROOM, M G. Identification and structure of the mating-type locus and

development of PCR-based markers for mating type in powdery mildew fungi. **Fungal Genetics and Biology**, v.48, p. 704-713. 2011.

COYLE, P.; PHILCOX, J. C.; CAREY, L. C.; ROFE, A. M. Metallothionein: the multipurpose protein. **Cell Mol Life Sci**. 59(4):627-47. 2002.

DEBUCHY, R., TURGEON, B.G., Mating-type structure, function, and evolution in Euscomycetes. In: Kues, U., Fischer, R. (Eds.), *The Mycota I Growth, Differentiation, and Sexuality*. Springer-Verlag, Berlin, Heidelberg, pp. 293–323. 2006.

DEBUCHY, R.; LECELLIER, V. B.; SILAR, P. Mating systems and sexual morphogenesis in Ascomycetes. **Cellular and Molecular Biology of filamentous fungi**. AMS Press, Washington, DC, p501-535, 2010.

DU, M.; SCHARDL, C. L.; NUCKLES, E. M.; VAILLACOURT, L. J. Using mating-type gene sequences for improved phylogenetic resolution of *Collectotrichum* species complexes. **Mycologia**, v97, n. 3, p. 641-658. 2005.

DYER, P. S.; FURNEAUX, P. A.; DOUHAN, G.; MURRAY, T. D. A Multiplex PCR Test for Determination of Mating Type Applied to the Plant Pathogens *Tapesia yallundae* and *Tapesia acuformis*. **Fungal Genetics and Biology**, v. 33, p. 173-180. 2001.

EFSA - EUROPEAN FOOD SAFETY AUTHORITY. Scientific Opinion on the risk of *Phyllosticta citricarpa* (*Guignardia citricarpa*) for the EU territory with identification and evaluation of risk reduction options. **EFSA Journal**, 2014; 12(2), Italy.

FU, R.; YIN, C.; DENG, Q.; LI, P.; ZHENG, A. Mating type loci analysis indicates that *Villosiclava virens*, the causal agent of false smut disease of rice, is a homothallic fungus. **Tropical Plant Pathology**, v. 39, n. 3, p. 203-209. 2014.

FUNDECITRUS. **Manual técnico sobre Pinta Preta**. Fundo Paulista de Defesa da Agricultura, 2008. 10p. (Boletim Técnico Edição especial). Disponível em < <http://www.fundecitrus.com.br/ImageBank/PageFlip/pageflip.aspx?idPage=71>>

GLIENKE, C.; PEREIRA, O.L.; STRINGARI, D.; FABRIS, J.; KAVA-CORDEIRO, V., GALLI-TERASAWA, L.; CUNNINGTON, J.; SHIVAS, R.G.; GROENEWALD, J.Z.; CROUS, P.W. Endophytic and pathogenic *Phyllosticta* species, with reference to those associated with Citrus Black Spot. **Persoonia: Molecular Phylogeny and Evolution of Fungi**, v. 26, p. 47-56, 2011.

GROENEWALD, M.; BARNES, I.; BRADSHAW, R. E.; BROWN, A. V.; DLAE, A.; GROENEWALD, J. Z.; LEWIS, K. J.; WINGFIELD, B. D.; WINGFIELD, M. J.; CROUS, P. W. Characterization and Distribution of Mating Type Genes in the Dothistroma Needle Blight Pathogens. **Mycology**, v. 97, n. 7. 2007.

HENNINGS, P. Fungi S. Paulenses IV a cl. Puttemans collecti. Hedwigia 48: 1 – 20. 1908.

KATOH, K.; TOH, H. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9, 286-298. 2008.

KOHMOTO, K.; YODER, O. C. Molecular Genetics of Host-Specific Toxins in Plant Disease. Japão: Proceedings of the 3rd Tottori International Symposium on Host-Specific Toxins. 24-29 Agosto, 1997.

KOTZÉ, J.M. Epidemiology and control of citrus black spot in South Africa. **Plant Disease**. V.65, p.945-950. 1981.

LI, W.; SULLIVAN, T. D.; WALTON, E.; AVERETTE, A. F.; SAKTHIKUMAR, S.; CUOMO, C. A.; KLEIN, B. S.; HEITMAN, J. Identification of the Mating-Type (*MAT*) Locus that controls sexual reproduction in *Blastomyces dermatidis*. **Eukaryotic Cell**, v. 12, n. 1, p. 109-117. 2013.

MABA, L. S. F. ***Phyllosticta citricarpa*: compatibilidade e mating types no ciclo sexual e aspectos da interação com o hospedeiro *Citrus sinensis***. 109 f. Tese - Programa de Pós – Graduação em Genética – Setor de Ciências Biológicas, Universidade Federal do Paraná, 2014.

O'DONNELL, K.; WARD, T. J.; GEISER, D. M.; KISTLER, H. C.; AOKI, T. Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. **Fungal Genetics and Biology**, v. 41, p. 600-623. 2004.

PACHECO, I. S. **Estudo comparativo do genoma e do transcriptoma de *Phyllosticta citricarpa* (*Guignardia citricarpa*) agente da Mancha Preta dos Citros e *Phyllosticta capitalensis***. 92 f. Dissertação – Instituto de Biociências da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de Botucatu, 2014.

PALMER, J. M.; KUBATOVA, A.; NOVAKOVA, A.; MINNIS, A. M.; KOLARIK, M.; LINDNER, D. L. Molecular Characterization of a Heterothallic Mating System in *Pseudogymnoascus destructans*, the Fungus Causing White-Nose Syndrome of Bats. **Genetics of Sex**. V.4. p.1755- 1763. 2014.

PÖGGELER, S. Mating-type genes for classical strain improvements of ascomycetes. **Applied Microbiology and Biotechnology**, v. 56, p. 589–601. 2001.

RUBINI, A.; BELFIORI, B.; RICCIONI, C.; TISSERANT, E.; ARCIONI, S.; MARTIN, F.; PAOLOCCI, F. Isolation and characterization of MAT genes in the symbiotic ascomycete *Tuber melanosporum*. **New Phytologist**, v. 189, p. 710-722. 2010.

SCHOCH, C. L.; CROUS, P. W.; GROENEWALD, J. Z.; BOEHM, E. W. A.; BURGESS, T. I.; *et al.* A class-wide phylogenetic assessment of *Dothideomycetes*. **Studies in Mycology**, v. 64, p1-15. 2009.

SINGH, G.; ASHBY, A. M. Cloning of the mating type loci from *Pyrenopeziza brassicae* reveals the presence of a novel mating type gene within a discomycete MAT 1-2 locus encoding a putative metallothionein-like protein. **Molecular Microbiology**. 30(4), 799-806. 1998.

SOLOVYEV, V.V. Statistical approaches in Eukaryotic gene prediction. In **Handbook of Statistical genetics** (eds. Balding D., Cannings C., Bishop M.), Wiley-Interscience; 3d edition, 1616 p. 2007.

TAMURA, K.; STECHER, G.; PETERSON, D.; FILIPSKI, A.; KUMAR, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30: 2725-2729. 2013.

TSUI, C. K. M.; DIGUISTINI, S.; WANG, Y.; FEAU, N.; DHILLON, B.; BOHLMANN, J.; HAMELIN, R.C. Unequal Recombination and Evolution of the Mating-Type (MAT) Loci in the Pathogenic Fungus *Grosmannia clavigera* and Relatives, *G3* (Bethesda), v.3, n.3, p.465-80, 2013.

TUCKER, S. L.; THORNTON, C. R.; TASKER, K.; JACOB, C.; GILES, G.; EGAN, M.; TALBOT, N. J. A fungal metallothionein is required for pathogenicity of *Magnaporthe grisea*. **The Plant Cell**, v. 16, p. 1575-1588. 2004.

TURGEON, B. G. Application of Mating Type Gene Technology to Problems in Fungal Biology. **Annu. Rev. Phytopathol.** V. 36, p. 115-137. 1998.

TURGEON, B. G.; YODER, O.C. Proposed Nomenclature for Mating Type Genes of Filamentous Ascomycetes. **Fungal Genetics and Biology**, v. 31, p. 1-5. New York. 2000.

YOKOYAMA, E.; ARAKAWA, M.; YAMAGISHI, K.; HARA, A. Phylogenetic and structural analyses of the mating-type loci in *Clavicipitaceae*. **FEMS Microbiology Letters**, v. 264, p. 182-191. 2006.



YUN, S. H.; ARIE, T.; KANEKO, I.; YODER, O.C.; TURGEON, B.G. Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. **Fungal Genetics and Biology**, v. 31, p. 7–20. 2000.

## 5. DISCUSSÃO GERAL

O estudo dos genes *mating type* é extremamente importante pois eles influenciam diretamente no ciclo da Mancha Preta dos Citros através do controle da reprodução sexual e produção dos ascósporos, responsáveis pela dispersão do fungo a longas distâncias. Neste trabalho, todas as sequências obtidas de cada um dos dez representantes de cada um dos genes *mating type* sequenciados foram idênticas, resultado compatível com estudos realizados com outros fungos, onde não foram encontradas variações intra-específicas significativas (TSUI *et al.*, 2013). É mostrado aqui que o fungo *Phyllosticta citricarpa* é heterotático, apresentando em seu genoma o idiomorfo *MAT1-1* (alfa box) ou *MAT1-2* (HMG box) e, portanto, requer isolados de ambos os *mating types* para completar seu ciclo sexual.

Um protocolo de PCR Multiplex *mating type* específico foi desenvolvido e este é capaz de identificar se as linhagens de *P. citricarpa* testadas possuem o idiomorfo *MAT1-1* ou *MAT1-2*, gerando amplificações de fragmentos com diferentes números de pares de base em uma mesma reação e facilmente identificáveis em gel de agarose. Esse teste pode ser utilizado em estudos de distribuição e/ou incidência de diferentes *mating types* e prever o risco de reprodução sexuada em determinada região. Além disso, este PCR multiplex é extremamente importante para a seleção de linhagens de *mating types* opostos e assim tentar a produção de ascos em laboratório para a realização de ensaios de indução de sintomas de MPC e controle da doença.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ARZANLOU, M.; CROUS, P. W.; ZWIERS, L. H. Evolutionary dynamics of mating-type loci of *Mycosphaerella* spp. occurring on banana. **Eukaryotic Cell**, v. 09, n. 01. 2009.

AZEVEDO, J. L.; ESPOSITO, E. **Fungos: uma introdução à biologia, bioquímica e biotecnologia**. 2 ed. Caxias do Sul: Educs, 638 p., 2010.

BAAYEN, R. P.; BONANTS, P. J. M.; VERKLEY, G.; CARROLL, G. C.; VAN DER Aa, H. A.; WEERDT, M de; VAN BROUWERSHAVEN, I. R.; SCHUTTE, G. C.; MACCHERONI JR, W.; GLIENKE DE BLANCO, C.; AZEVEDO, J. L. Nonpathogenic Isolates of the Citrus Black Spot Fungus, *Guignardia citricarpa*, Identified as a Cosmopolitan Endophyte of Woody Plants, *G. mangiferae* (*Phyllosticta capitalensis*). **Ecology and Population Biology**. V. 92, n. 5. 2002.

BALDASSARI, R. B.; REIS, R. F.; GOES, A. A new method for inoculation of fruit with *Guignardia citricarpa*, the causal agent of citrus black spot. **European Journal of Plant Pathology**, v. 123, n. 1, p. 1–4, 2008.

BIHON, W.; WINGFIELD, M. J.; SLIPPERS, B.; DUONG, T. A.; WINGFIELD, B. D. MAT gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. **Fungal Genetics and Biology**, v. 62, p. 55-61. 2014.

BÖHM, J.; HOFF, B.; O'GORMAN, C. M.; WOLFERS, S.; KLIX, V.; BINGER, D.; ZADRA, I.; KÜRNSTEINER, H.; PÖGGELER, S.; DYER, P. S.; KÜCK, U. Sexual reproduction and mating-type-mediated strain development in the penicillin-producing fungus *Penicillium chrysogenum*. **PNAS**, v.110, n.4, p. 1476-1481. 2012.

CHEN, R. S.; CHU, C.; CHENG, C. W.; CHEN, W. Y.; TSAY, J. G. Differentiation of two powdery mildews of sunflower (*Helianthus annuus*) by a PCR-mediated method based on ITS sequences. **Eur J Plant Pathol**. V. 121, p. 1-8. 2008.

CHITRAMPALAM, P.; INDERBITZIN, P.; MARUTHACHALAM, K.; WU, B. M.; SUBBARAO, K. V. The *Sclerotinia sclerotiorum* Mating Type Locus (*MAT*) Contains a 3.6-kb Region That Is Inverted in Every Meiotic Generation. **Plos One**, v. 8, n. 2. 2013.

CITRUSBR. Brasil é o maior produtor mundial de laranja. 2013. Disponível em <<http://www.uagro.com.br/editorias/agricultura/laranja/2013/08/07/brasil-e-o-maior-produtor-mundial-de-laranja.html>> Acesso: Setembro, 2015.

COPPIN, E.; DEBUCHY, R.; ARNAISE, S.; PICARD, M. Mating types and sexual development in filamentous ascomycetes. **Microbiology and Molecular Biology Reviews**, v. 61, n. 4, p. 411-428. 1997.

CÔTÉ, M. J.; TARDIF, M. C.; MELDRUN, A. J. Identification of *Monilinia fructigena*, *M. fructicola*, *M. laxa*, and *Monilia polystroma* on Inoculated and Naturally Infected Fruit Using Multiplex PCR. **Plant. Dis.** V. 88, p. 1219-1225. 2004.

DEBUCHY, R.; LECELLIER, V. B.; SILAR, P. Mating systems and sexual morphogenesis in Ascomycetes. **Celular and Molecular Biology of filamentous fungi**. AMS Press, Washington, DC, p501-535, 2010.

DEMEKE, T.; CLEAR, R. M.; PATRICK, S. K.; GABA, D. Species-specific PCR-based assays for the detection of *Fusarium* species and a comparison with the whole seed agar plate method and trichothecene analysis. **International Journal of Food Microbiology**. V. 103, p. 271-284. 2005.

DYER, P. S.; FURNEAUX, P. A.; DOUHAN, G.; MURRAY, T. D. A Multiplex PCR Test for Determination of Mating Type Applied to the Plant Pathogens *Tapesia yallundae* and *Tapesia acuformis*. **Fungal Genetics and Biology**, v. 33, p. 173-180. 2001.

EFSA - EUROPEAN FOOD SAFETY AUTHORITY. Scientific Opinion on the risk of *Phyllosticta citricarpa* (*Guignardia citricarpa*) for the EU territory with identification and evaluation of risk reduction options. **EFSA Journal**, 2014; 12(2), Italy.

ELNIFRO, E. M.; ASHSHI, A. M.; COOPER, R. J.; KLAPPER, P. E. Multiplex PCR: Optimization and application in diagnostic virology. **Clinic Microbiology Reviews**, v. 13, n. 4, p. 559-570. 2000.

FAOSTAT – Food and Agriculture Organization of the United Nation. Disponível em <<http://faostat3.fao.org/home/E>> Acesso: Setembro, 2015.

FEICHTENBERGER, E. Mancha preta ou pinta preta dos citros. **Laranja & Cia.** v. 43, p. 10-11, 1996a.

FUNDECITRUS. **Manual técnico sobre Pinta Preta. Fundo Paulista de Defesa da Agricultura**, 2008. 10p. (Boletim Técnico, Edição especial). Disponível em <<http://www.fundecitrus.com.br/ImageBank/PageFlip/pageflip.aspx?idPage=71>>. Acesso: Maio 2015.

FUNDECITRUS. Fundecitrus estima safra de 278,993 milhões de caixas de laranja no cinturão citrícola. Disponível em: <<http://www.canalrural.com.br/noticias/citrus/fundecitrus-estima-safra-278993-milhoes-caixas-laranja-cinturao-citricola-56544>> Acesso: Maio 2015.

GLIENKE, C.; PEREIRA, O.L.; STRINGARI, D.; FABRIS, J.; KAVA-CORDEIRO, V., GALLI-TERASAWA, L.; CUNNINGTON, J.; SHIVAS, R.G.; GROENEWALD, J.Z.; CROUS, P.W. Endophytic and pathogenic *Phyllosticta* species, with reference to those associated with Citrus Black Spot. **Persoonia: Molecular Phylogeny and Evolution of Fungi**, v. 26, p. 47-56, 2011.

GROENEWALD, M.; BARNES, I.; BRADSHAW, R. E.; BROWN, A. V.; DLAE, A.; GROENEWALD, J. Z.; LEWIS, K. J.; WINGFIELD, B. D.; WINGFIELD, M. J.;

CROUS, P. W. Characterization and Distribution of Mating Type Genes in the Dothistroma Needle Blight Pathogens. **Mycology**, v. 97, n. 7. 2007.

HAWKSWORTH, D. L.; CROUS, P. W.; REDHEAD, S. A; REYNOLDS, D. R.; SAMSON, R. A.; SEIFERT, K. A.; TAYLOR, J. W.; GLIENKE, C.; *et al.* The Amsterdam Declaration on Fungal Nomenclature. **IMA Fungus**, v. 2, p. 105-112, 2011.

HEITMAN, J.; SUN, S.; JAMES, T. Y. Evolution of fungal sexual reproduction. **Mycologia**, v. 105, n. 01, p. 1-27. 2013.

HENDOLIN, P. H.; MARKKANEN, A.; YLIKOSKI, J.; WAHLFORS, J. J. Use of Multiplex PCR for Simultaneous Detection of Four Bacterial Species in Middle Ear Effusions. **Journal of Clinical Microbiology**, v. 35, n. 11, p. 2854-2858. 1997.

HENEGARIU, O.; HEEREMA, N. A.; DLOUHY, S. R.; VANCE, G. H.; VOGT, P. H. Multiplex PCR: Critical parameters and step-by-step protocol. **BioTechniques**, v. 23, n. 3, p. 504-511. 1997.

HOANG, M.L.; TAN, F.J.; LAI, D. C.; CELNIKER, S. E.; HOSKINS, R. A.; DUNHAN, M. J.; ZHENG, Y.; KOSHLAND, D. Competitive repair by naturally dispersed repetitive DNA during non-allelic homologous recombination. **PLoS Genet** 6: e1001228. 2010.

HUANG, C. S. *et al.* Leaf infection with citrus black spot and perithecial development in relation to ascospore discharge of *Guignardia citricarpa* Kiely. **Journal of Taiwan Agricultural Research**, v. 21, n. 4, p. 256-263, 1972.

IBGE – Estatística da Produção Agrícola. Disponível em: < [http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/estProdAgr\\_201309.pdf](http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/estProdAgr_201309.pdf)> Acesso: Agosto 2015.

KLIX, V.; NOWROUSIAN, M.; RINGELBERG, C.; LOROS, J. J.; DUNLAP, J. C.; PÖGELLER, S. Functional characterization of *MAT-1-1*-specific Mating-Type genes in the homothallic Ascomycete *Sordaria macrospora* provides new insights into essential and nonessential sexual regulators. **Eukaryotic Cell**, v. 09, n. 06, p. 894-905. 2010.

KOTZÉ, J.M. Epidemiology and control of citrus black spot in South Africa. **Plant Disease**. V.65, p.945-950. 1981.

KRONSTAD, J. Self-Fertility: The Genetics of Sex in Lonely Fungi. **Current Biology**, v. 17, p. 844, 2007.

LI, W.; SULLIVAN, T. D.; WALTON, E.; AVERETTE, A. F.; SAKTHIKUMAR, S.; CUOMO, C. A.; KLEIN, B. S.; HEITMAN, J. Identification of the Mating-Type (*MAT*) Locus that controls sexual reproduction in *Blastomyces dermatidis*. **Eukaryotic Cell**, v. 12, n. 1, p. 109-117. 2013.

MABA, L. S. F. ***Phyllosticta citricarpa*: compatibilidade e mating types no ciclo sexual e aspectos da interação com o hospedeiro *Citrus sinensis***. 109 f. Tese - Programa de Pós – Graduação em Genética – Setor de Ciências Biológicas, Universidade Federal do Paraná, 2014.

McONIE, K.C. Source inoculum of *Guignardia citricarpa*, the citrus black spot pathogen. **Phytopathology** 54, 64-67, 1964a.

METZENBERG, R. L., and GLASS, N. L. 1990. Mating type and mating strategies in *Neurospora*. **BioEssays** 12:53-59.

NEVES, M. F.; TROMBIN, V. G.; MILAN, P.; LOPES, F. F.; CRESSONI, F.; KALAKI, R. **O Retrato da Citricultura Brasileira**. São Paulo: Maskestrat., 138 p., 2010.

OTTO, S. P. The Evolutionary Enigma of Sex. **American Society of Naturalists**, v. 174, p. S1-S14. 2009.

PAOLETTI, M.; RYDHOLM, C.; SCHWIER, E. U.; ANDERSON, M. J.; SZAKACS, G.; LUTZONI, F.; DEBEAUPUIS, J. P.; LATGÉ, J. P.; DENNING, D. W.; DYER, P. S. Evidence for Sexuality in the Opportunistic Fungal Pathogen *Aspergillus fumigatus*. **Current Biology**, v. 15, p. 1242-1248. 2005.

PAOLETTI, M.; SEYMOUR, F. A.; ALCOCER, M. J. C.; KAUR, N.; CALVO, A. M.; ARCHER, D. B.; DYER, P. S. Mating Type and the Genetic Basis of Self-Fertility in the Model Fungus *Aspergillus nidulans*. **Current Biology**, v. 17, p. 1384-1389. 2007.

PÖGGELER, S.; NOWROUSIAN, M.; KÜCK, U. Fruiting-Body Development in Ascomycetes. The Mycota I Growth, Differentiation and Sexuality. Springer-Verlag Berlin Heidelberg: Kues/Fischer (Eds.) 2006.

POIREL, L.; WALSH, T. R.; CUVILLIER, V.; NORDMANN, P. Multiplex PCR for detection of acquired carbapenemase genes. **Diagnostic Microbiology and Infectious Disease**, v. 70, p. 119-123. 2011.

REIS, R. F.; TIMMER, L. W.; GOES, A. Effect of temperature, leaf wetness, and rainfall on the production of *Guignardia citricarpa* ascospores and on back spot severity on sweet orange. **Fitopatol. bras.**, v. 31, n. 1, 2006.

RUBINI, A.; BELFIORI, B.; RICCIONI, C.; TISSERANT, E.; ARCIONI, S.; MARTIN, F.; PAOLOCCI, F. Isolation and characterization of MAT genes in the symbiotic ascomycete *Tuber melanosporum*. **New Phytologist**, v. 189, p. 710-722. 2010.

SLIPPERS, B.; BOISSIN, E.; PHILLIPS, A. J. L.; GROENEWALD, J. Z.; LOMBARD, L.; WINGFIELD, M. J.; POSTMA, A.; BURGESS, T.; CROUS, P. W. Phylogenetic lineages in the Botryosphaerales: a systematic and evolutionary framework. **Studies in Mycology**, v. 76, p. 31-49. 2013.



SPOSITO, M. B., AMORIM, L., BASSANEZI, R. B., FILHO, A. B. AND HAU, B. Spatial pattern of black spot incidence within citrus trees related to disease severity and pathogen dispersal. **Plant Pathology**, v. 57, p. 103–108, 2008.

SUTTON, B.C. & WATERSTON, J.M. *Guignardia citricarpa* Descriptions of pathogenic fungi and bacteria. **Surrey**, England, Kew: Commonwealth Mycological Institute, n.85, 2p., 1966.

TALY, V.; PEKIN, D.; BENHAIN, L.; KOTSOPOULOS, S. K.; CORRE, D. L.; LI, X.; ATOCHIN, I.; LINK, D. R.; GRIFFITHS, A. D.; PALLIER, K.; BLONS, H.; BOUCHÉ, O.; LANDI, B.; HUTCHISON, J. B.; LAURENT-PUIG, P. Multiplex Picodroplet Digital PCR to Detect *KRAS* Mutations in Circulating DNA from the Plasma of Colorectal Cancer Patients. **Clinical Chemistry**, v. 59, n. 12, p. 1722-1731. 2013.

TANIUCHI, M.; VERWEIJ, J. J.; NOOR, Z.; SOBUZ, S. U.; LIESHOUT, L. V.; PETRI JR, W. A.; HAQUE, R.; HOUP, E. R. High Throughput Multiplex PCR and Probe-based Detection with Luminex Beads for Seven Intestinal Parasites. **Am. J. Trop. Med. Hyg**, v. 84, n. 2, p. 332-337. 2011.

TIMOSSI, A. J.; GOES, A. de; KUPPER, K. C.; BALDASSARI, R. B.; REIS, R. F. dos. Influência da Temperatura e da Luminosidade no Desenvolvimento de *Guignardia citricarpa*, Agente Causal da Mancha Preta dos Frutos Cítricos. **Fitopatologia Brasileira**, v. 28, p. 489-494, 2003.

TSUI, C. K. M.; DIGUISTINI, S.; WANG, Y.; FEAU, N.; DHILLON, B.; BOHLMANN, J.; HAMELIN, R. C. Unequal Recombination and Evolution of the Mating-Type (MAT) Loci in the Pathogenic Fungus *Grosmannia clavigera* and Relatives, *G3* (Bethesda), v.3, n.3, p.465-80, 2013.

TURGEON, B. G., BOHLMANN, H., CIUFFETTI, L. M., CHRISTIANSEN, S. K., YANG, G., SCHAFER, W., and YODER, O. C. Cloning and analysis of the mating type genes from *Cochliobolus heterostrophus*. **Mol. Gen. Genet.** 238:270-284. 1993.

TURGEON, B. G.; YODER, O.C. Proposed Nomenclature for Mating Type Genes of Filamentous Ascomycetes. **Fungal Genetics and Biology**, v. 31, p. 1-5. New York. 2000.

YOKOYAMA, E.; ARAKAWA M.; YAMAGISHI, K.; HARA, A. Phylogenetic and structural analyses of the *mating type* loci *Clavicipitaceae*. **FEMS Microbiol Lett.** V. 264, p. 182-191. 2006.

YUN, S. H.; ARIE, T.; KANEKO, I.; YODER, O.C.; TURGEON, B.G. Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. **Fungal Genetics and Biology**, v. 31, p. 7–20. 2000.

WANG, X.; CHEN, G.; HUANG, F.; ZHANG, J.; HYDE, K. D.; LI, H. *Phyllosticta* species associated with citrus diseases in China. **Fungal Diversity**, v. 52, p. 209-224. 2012.

WIKKEE, S.; LOMBARD, L.; NAKASHIMA, C.; MOTOHASHI, K.; CHEEWANGKON, R.; MCKENZIE, E. H. C.; HYDE, K. D.; CROUS, P. W. A phylogenetic re-evaluation of *Phyllosticta* (*Botryosphaerales*). **Studies in Mycology**, v. 76, p. 1-29. 2013.

WOLCOTT, M. J. Advances in nucleic acid-based detection methods. **Clin. Microbiol. Rev.** 5:370–386. 1992.

WULANDARI, N. F.; TO-ANUN, C.; HYDE, K. D.; DUONG, L. M.; GRUYTER, DE J.; MEFFERT, J. P.; GROENEWALD, J. Z.; CROUS, P. W. *Phyllosticta citrasiana* sp. nov., the cause of Citrus tan spot of *Citrus maxima* in Asia. **Fungal Diversity**. Holanda, v. 34, p. 23-39, 2009.

ZAFFARANO, P. L.; DUÒ, A.; GRÜNIG, C. R. Characterization of the mating type (MAT) locus in the *Phialocephala fortinii* s.l. – *Acephala applanata* species complex. **Fungal Genetics and Biology**. V. 47, p. 761-772. 2010.

ZHENG, Z.; LIEBERS, M.; ZHELYAZKOVA, B.; CAO, Y.; PANDITI, D.; LYNCH, K. D.; CHEN, J.; ROBINSON, H. E.; SHIM, H. S.; CHMIELECKI, J.; PAO, W.; ENGELMAN, J. A.; IAFRATE, A. J.; LE, L. P. Anchored multiplex PCR for targeted next-generation sequencing. **Nature Medicine**, v. 20, n. 12. 2014.

## ANEXOS

### ANEXO1: ESTRUTURA E PREDIÇÃO DOS GENES DO LOCUS MAT1 EM *Phyllosticta citricarpa*

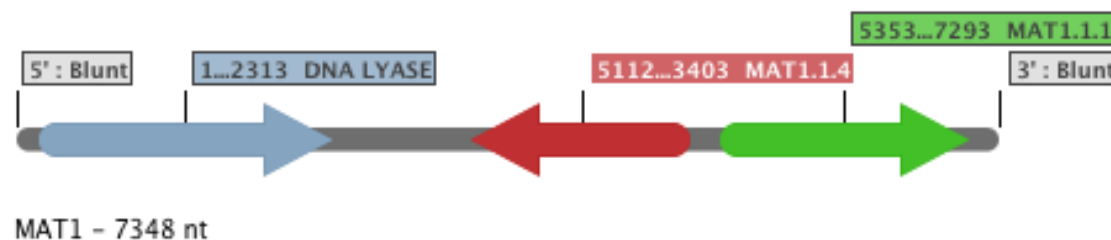


FIGURA 1: DNA LYASE + MAT1.1.4 + MAT1.1.1 DE *P. CITRICARPA*

> DNA LYASE + MAT1.1.4 + MAT1.1.1 DE *P. CITRICARPA*

```
ATGCCGCTCCGGATAACCTCGTGGAACGTCAATGGCATCCGCAATCCTTTTGGCTACCAGCCATGGCGCGACAAGCGGTCCTTCGACGC
CATGTTCGACCTCCTCGAGGCCGACATTGTGCTCATGCAAGAGCTCAAGATTCAGCGCAAGGACCTCACCGACGACATGGTCCTGGTTC
CCGGCTGGGATTGCTACTTTAGTCTGCCAAAGCACAAAAAAGGTGCGCGTGCTGCCGCTCTGATGTGACGGCTGCTGACGGCTTCAGGA
TACTCTGGCGTAGCGATATACTCGCCAATCTGCCTGCGTGCCCATTCGCGCTGAAGAGGGCGTCCTGGGCGTCCTCTGCCCTCCAGG
ATCCTCCACGTCTTACCGCAATCTGCCCGCCGACCAAGCCATTGGCGGATATCCAACCGACGAACAAATTGCCGCCGCCGCGCTCGACC
CTGCTGCGTTGGATGCCGAAGGGCGCTGCGTCGCTCTCGAGTTTCCTGCCTTTGTTCATTCTCGGCGTCTACAGCCCGGCGAAGTTCGAAT
GGAATGAGAGACGACTTCCGCTACGGGTTCTCTGCGCCTTGGACGCGCGCATACTGTAATTTGACGGCGTTGGGCAAGAATATCATTCT
GACTGGCGACCTCAATGTCTCGCGGCAAGAACTGGACACGGCGAATGCGGAGGAGACGCTCAAGAAGGAGGGCATCACGCACGAAGAGT
ACGTGTCGTCGCCAACAGGCGCATGTTTAACCAGCTGCTGGAAAATGGCGAGGTTGTGGGTAGTCGAGACGAAGGACGGGAGCAGCCG
GTGCTGTGGGACACGTGCCGAGAGTTTCATCCTACGCGCAAGGGCATGTACACGCATTGGGAGCAAAAGATTAATGCGAGACCGGGAAA
```

TTTTGGCTCCAGGATCGACTATGTCCTTTGCAGTATTGCAATGAAGGATTGGATCTGCGACGCCAACATTCAGGAGGGCCTCATGGTTA  
GTTGATGCGCTTCGATGTCTTGAACGACTGCTGACTTTTGTTCAGGGCTCTGATCATTGCCCAGTATATGCCATGCTCAAGGATCGTGT  
CACGCTCAATGGCGTCGAGTATGATTTGCCAGAGCTGATGAATCCTCCCGGCACGTTTCAAGACGGCCAAAGGAAAATCGAATGGAGCT  
TCAAGAGTTTGCCCGCGCTCTGTGGACGACTGCTTCCAGAGTTTGACCGTCGAAGGAGCATCAAGGACATGTTTACGCGCAAGAAACCG  
GCCTCTGCTGCTTCTCCGGGTGCATCTATTTCTGTTGCAGCAACTCATTGAGCAACTTCTTCAAATGGTTCGCCCCGGTGTTCAGCATCC  
CACTGAGCCCCGAACCAAATGCTGCATCTGCAATATCGAGTGAGCCACAGCAACCGTCAGAACAAAAGGCGCCGTTTCGCGGAGCCTACTA  
AAAGCCCAGAGAAGAAACGAAATGCTGCTGGAGGTCCTACCGCCAAATCCACAAAACGGCAGAAGCTTGGCCCTGCTTCGTCTAAATCG  
AATCCCGCAAAGGGTCAACAGTCATTGACGGGATTCCCTCAAGCCAGCCGCGAGGGCGACGAACAATGGTAGCGCTGTAAACGAGCCTGG  
TTCGAGGAAGACCATGATAGAAAAGGAACGACAAACAACCTCCCGTCAATCTCTACTCCGACCCTGGAAGCCCGTCTCCCAAGTATACGT  
CGTCTCAGGGCGCTGCGTCGCAAATGACAGAAGACCTCAATGGGACGACCGCAAACCTCTCTAAGAACCAGTGAAAAGACGGCAGCCTCA  
GGGGACTACCGCTCTACCGCAATCGAATCAGCCTCCAACTCGTCAGTGGCACTGGAAGATTTGAGCGATATCCAGACTCGCGAGAAAC  
GTGGTCGCGGCTCTTTGCCAAACGCGATCCACCGCGCTGCGAAGACCATGGCGAGTTTTTGCATCAGCCTCACGACCAAGAAGAGCGGCG  
TCAACTGTGGGAGAGCATTTTTGGATATGTCCGAGGTAAGCACTTCGCCCAGTTGTGGCAAGCATGATTAGTTTGTTCGATTGTCAATGT  
CATACTGGCTTTTACCATCTTGGTGCCACGCATTTACGATTACATCTTTCCAATGTGCAGCTTCTTGATTGGAATGCTGCATGACCTTT  
CCCCTCCCGCCATTCATCTTTTCCAAGACGCAATCAAGTTGAGGCTTTTGTATGCTGACCCTCCTTGCTTTTTTGCAGACCAATCGGCCC  
CAGCGGGTCCAAAGAAAAAGGGACAGAATGGCGCTGCCCCACCTTCATCTGGGCCAGCGATTGGAACGGCTCCACCGCAGCACCATGAA  
GGCGAGTGTCCAAAGGCACGACGGCAGATTTATGTGTGGAAGAGCGCAGCAGAGAATGATTTCTCGCGCATATCGGACCAACCCACCGG  
GTGACCCATCACATCGACGCCTCCGTAGCTGAAGTCTGCCGATCCCTGAAGGTTGCCCACTACAGCCCTGTCGTGGCAGGTTCTTCCCG  
AACTGTCTTGCCCTTGCTAGACCGTCAGACAACTTGCCCTACGATTTATGAAACATGCCCTCTCCATGGTCCGGTTGTCCATGCGTTTCT  
CTGTCTGTCAAACCTGCGGGTAGGATCCCGTGTATCAACAAGTGTACAATTGACGGGGCCGAGCTCGGGGTGTGTTTGTTCCTCTGTTT  
CCATGTCTCGTTTTTCCGTCTTTTCCAGAGAGGTGCGACAGGCAACAGGGGAAGAGAGACAGGCGCATGAATCAAGACAACGGCGCCACGAG  
ACAGGGCAAGACAGAAGAACTGGTGGCACACGTGGCTACGAGACAAGCCGTGTTGCTGGCACGGCTACGGCATGGCAGACCATGGCGA  
AATGGGTTTCCGTCCACAACGCGCCCTGCCACGGCGTGCATCTCTGCCGCTCGTCCAGTGCACCGGCCACCAGCTAGCTGATCGACTTA  
CACAGTGACGTATCCAGCGTTTGCATTGCGCAGATTGGATTCTGAGAGTTACGACTGCCGCCTTCACCGCCTTCACCGCCGCCATACCA  
AACATTGATGTCGCGATGTTTCGTAAATTCGTAGACTGCAAAACAAATCGCTACGTATGTAGCTCTTTCCCCATGCGCCCCAACGCCGTC  
GGTGCAGCCGTCATATGCCAGCCCTGCCACGCGCTTGCTCGTAGCCCTCCCGCCGCTCGCGGGCGCCGCTTGCGTGGAAGAAAGCGAATG

CAAGAGAAGAGATGGCTAGTTTGTGTTTTTGGCCGCCTTCTCACGTCGTCCTTCTCCTCCGCTGCTCGCGCCCAGCAGCATCCGCGTCC  
CGATATCCCTCCTTCAAACGAAGCCATCCAACCTATCTCCGCGTCGACCTCGCAACGTACATAGCCATCCATCCCACACACCATCACCA  
CGCCACACCTCCACCTCCATCTACGCAGCAGCGCCAGCCGCCGAGCTCTCGCCCTGTTGCTGCCCTGCTGCTGCTGCGGGCCCTCCCC  
CTGCACAACCTGCATCTTCCGCCTTCGTTCCCTTGTCCGCCTTTTCCTTGCGGTCCGCGGGCGTCCTTGTTCTCTGAGGAGGAGCGCTTGA  
CGAAGCGCTGCGTCAGCACGGCGTCGAAGGCGTCGTCGTCGTCGTCGTCGGTCGGGCGGGCTGGAAGGCCTCGTCGCGGCTGCTGACGCGC  
GGCATCAGCGCGCCCAGGCGCCGGTCGAAGCGCGATACGTCGCGGAAGGTGTGGGTGCGGTGGTTCGAGCTGTGAGTGGGAGAGGGGGGT  
GTTAGTGGCGTGTTTGTGTGGGTGGTGGGTGGGAGGGTTGGGGATGGGAGTATGGCGTCTGTGGACGTACGTACCAGGTTCCATGGGC  
TGGGCTCGAATGTCATGGGCTCGAAAGTGCGCAGGCCAAAGTTGAGGCGGTGCATGCTTGATGGCTCCACTCGATGCGCTGCGTTGCG  
TCGTTAGCTCGTTTGCCTCTTCGTCGTCCATCATCGGCATCAGCATCTGCTGCCTGTGTGCTCCTCCCCTTCTCTTTCTCTTCTGCTAT  
GCCGAGACGCGGTTCAGAGGAGCCAAAACGTACCTTGTGCCAAAGCGACGCCTGTGCTGGCCGCAGCTGCAACGATGTTAGCGGGAGCTG  
AGAGAAGAGTAGGAAATGTCCACACCAAGTCGATTGCCTGAATCATCGCATAGAGCCGGTCGTCCCAAACTGAAGAATCTTAACGCCG  
CCCCGCGGAATCAGTTGATACGCATCGCGCTTCTGATGCAGCACGCGCCGCATCTGAAGCAGCAGGATACTAGCACCCGGCGCATTTCTT  
ATCCAAAACTGACTCCAGCGCGGCAAATCCCTGACAAACCCATTCTTATTCCTATTGACGCGCTCGCCTTCCTTGATCTTGGCCACCA  
CGTCGCGGACCTTGCGCTTGAATTCGTCAAGTCGATGAGGTCGGTCGACAGTTTGTGGAGCGCGGCGCAGAAGCGGCGGCGGTGGGG  
GCGTCCTCGATGAGGCGGGCGGAGCGCAGCACGCGGGCGGCACTCGTCGGCCAGCTTCTCGATCGAGATGCGCAGGTGCTGCTCGTCGG  
TTGCCCCAAGTCGAGGATGTATTGGTTTCGCAGCGTCTCGAAGTCTTGGATCAGGTCGGCCAGCTGTGCGTCGGAGCCGAGGATCAGAG  
TGGCTGCGGTTTCGTCAGCGCGGTTCGATGTGGGAATTGGCGCGTGTGCTTGCCTTGTGAGCCTGTGAATCCGGAGCCCTTGAGGCCTGCT  
TCTGGATCTTCAATTCGGAGGCGCGGTTTTTTTCTTGGCGCGCTCGAGGAGAGGCGCTTGTGGGTTCAGAGGTCCATGGCTTGTGGGGT  
GTGATGGGCAAAGGTTAATCTCTAGGTGCCTGGCTCTATCGGAGGAAGAAAGGAGTCAGCGTTTTTCGGAGCATCGTCTTCGTTGGTCAA  
AACTTGCGGGAAGATAATGCGCGCGTGGTATGTAGGCGAATGCGGAAATGTTGCCAAGACCGTCAAAGATGTATTTGTTGCAGACGCAG  
CATGTAGGTTGATGCAGATGTTGAGTTGTCAGAAGATTGCGCGAGGAAGCACGCTCCATGATGTCTAGGCTAAGAGTCTCACAGGAAAT  
AGCCGGCTGAGAAGGGCGGTCTGTGCTCTATCGAACATTTCCATGCCCTGTGATGTCGTCACCAAGGTGCTGGTGTGAAGTGAAGTGA  
TGCCCTTTGTTGTGACCAGCCATCAACCTTTCGGGGCACAAATGAAATTTACTCCACTCTCCAGCTTACATCCTCCATCCGCATCCTCCT  
CAAAACAATCTTCTTTGTCTCAGTTCACCCGAAACACCAGCTACCTCAACCATCGTTGCTTTCCGACTCTTCCTCATAGCCAACTTGTT  
CTTCGGCTCAAAATGGCTGTGCTGCAAACCTTCCTCCCCACGTCAAGTCATACCTGGCCGCATGCTCTCAAGCCCAAATCAATCAGCT  
TGTTTCATCGCTCAAAGGCGTCGAAGGCAAGAGGACAGTCGCCGCTGTGGTTGCCGCGGGCGACCTACAACCTCGTGGCTCGCACTCGA

CGCTCTCTTCCTCCAAGTCAACTTCCACCCAAGCGACAACCTCCGTCAGGCTCACGGTCGCCTTACATGACTCCCGCATTCCTCGACAAAG  
AAGACCTCCTCAAGCGCGCAGGCAAAGACCCTGCCAACTTCGCAGCCAGATAGAATCTCTGCAAACATGACCAATGTTCTCACGTTGC  
GCCAGACAGTTCGTCCAGTCGAGCTACATCTCAGACAGGCGCTACGCAGACAAGCCCTCTACGAGCTCTCAACTCTTTCATGGCTTTTC  
GAGGTTGGTGATCTGATTCGGAATTGCCGCCAAGACTTAACTAACAAGCGGTAGTTTACTACAATCCGATCTTCCACTCTTTTGAGCAG  
AAGGAGCGCTCAGGCTTCATGAAGATAATGTGGGCCGAGGAGGAAACCAAAGCGAAGTGGAGCATCATTGCCAAGGCATGGTCAATCAT  
CCGGGATGAAGTGGGCAAAGATTCTGCTCCGCTTGACAGATTCCTTGAGCTTGCCCTGCCCTCACATTGACATCCTCGCTCGAGATGCAT  
ACTTGTCGACCTTGGGTTGGGAGCTTGCGACGGACAATGGCGACACGTCTCTTAGACGGCAATTCCTGCCCAACTTCAAGGATTTTCCT  
GCCAGCTTCCGCACAACCTACAAAGTCAGTTCAGGACGTCAATTGACTTCGTCCGCAACTTGGGATACGGTTTACCTACCAATGACACGCC  
TCCAGCTTCAAGCAACATCTCGACCTCTCATCTGACGATGGCTGCACAGCCTGTTTCATCATTCCACCGCGACCACAGGCCCTGGATTCC  
TTCCCTCTTCGATATTCCACTCCACAACTCCAGCTCCAACGCCAACTTTCAGTCTCTTGACGCTTCCAATCTTGGAGGCCCCCATGCG  
TCCAAC TTCCCTGAGTCTGCTATCAGCAGCACGACGCAGACAACCTCAATGGCAGCCTCGACTACGCTATCCCACGCTGGTGACAATGC  
TTCAGCTCCAATTGAACAATCTGCACAGGAAAATGCAATGGATCTTGGCAGTGACGCGCAGCAATCAGCGGCAGGCACCAATCCTAGTT  
ATGCGGTCAGCCAGCAGCAGGCAAGCAATGCAACAAACGCACCACCTCCTCAAACAAGCAGCGTCAATGCAGCTTCTGGGCTTACGTCC  
GGTAGCAATGTGGCAGCGCAATCCAATGACCAGGCTCTGATGGCCATCATCCACCAACATGCTGCTGCTGCCAATGCTGGTAATGGAGG  
CCTGCTCACTGGAGGCCCCATGACTCACCAGCAGATGTTGACTTACATGTATGGGGCCAACGTACAACCAGCCCACCCCGACATGTACT  
CACCACACGTGGACGATACGATTGCGAATCCCGACCCCCAGAGTTGCGTGATGCAGCGAGCCAGCCTCGCCGATCCAGATCTTCTTGCC  
ATCTCCATCAACGCGGAATACCCTCATGAAATCCTGAACTATTACCCTGCTTCAACGCGCAATTCAGGAGAATCGATGGCAAGCTCGGT  
TGTGCGCCCGACCTGCGACGAGGACACGCAGACCGCCTGGCTGGACTGCGCCCTACAGATCATGGACACATCACCATCATCTTCACCAA  
CATTTCTTATGCCGAGTTCTTCAATCCTGACCATGAGGGCTTGAATGTCTTCGACATTGACGGAATTCCTGATTTTACTGCCGTTGAC  
AACACCGACTTCCAGTTCAGGAGATGTCCTTGCCCTTTGGTCACCAGCCACTGGGATGATGAGTCCGATGGTCCAGAGCAGTGAAAAGA  
GGACAGTCGGTCGGATTTGAAGTGTCATGAATATCCAGGAGTTAGGCTCG

Predicted genes for sequence number 1 on both strands

start gene g1 = DNA LYASE

```
unnamed-1AUGUSTUS gene 1      2313 0.73 +      .      g1
unnamed-1AUGUSTUS transcript    1      2313 0.41 +      .      g1.t1
unnamed-1AUGUSTUS start_codon   1      3      .      +      0      transcript_id "g1.t1"; gene_id
"g1";
unnamed-1AUGUSTUS initial      1      220 0.94 +      0      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS internal    266 975 0.54 +      2      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS internal    1025 1992 0.77 +      0      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS terminal     2214 2313 0.82 +      1      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS intron      221 265 0.54 +      .      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS intron      976 1024 0.89 +      .      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS intron      1993 2213 0.84 +      .      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 1        220 0.94 +      0      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 266 975 0.54 +      2      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 1025 1992 0.77 +      0      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 2214 2310 0.82 +      1      transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS stop_codon    2311 2313 .      +      0      transcript_id "g1.t1"; gene_id
"g1";
```

coding sequence =

```
[atgccgctccggataacctcgtggaacgtcaatggcatccgcaatccttttggctaccagccatggcgcgacaagcggtccttcgacg
ccatgttcgacctcctcgaggccgacattgtcggtcatgcaagagctcaagattcagcgcaaggacctcaccgacgacatgggtcctgggt
cccggtggtgattgctacttttagtctgccaaagcacaaaaaaggatactctggcgtagcgatatacactcgccaatctgcctgcgtgcc
cattcgcgctgaagagggcggtcctggggcggtcctctgccctccaggatcctccacgtcctaccgcaatctgcccgcggaccaagccattg
gcggatatccaaccgacgaacaaattgccgcccggcggtcgaccctgctgcgttggtatgccgaagggcgctgcgtcgctctcgagttt
cctgcctttgtcattctcggcggtctacagcccggcggaactcgaatggaatgagagacgacttccgctacggggttcctctgcgccttgga
```



cgcgcgcatacgtaatttgacggcggttgggcaagaatatcattctgactggcgacctcaatgtctcgcggaagaactggacacggcgga  
atgcggaggagacgctcaagaaggagggcatcacgcacgaagagtacgtgtcgcccaacaggcgcatgtttaaccagctgctggaa  
aatggcgaggttgtgggtagtcgagacgaaggacgggagcagccggtgctgtgggacacgtgccgagagtttcatcctacgcgcaaggg  
catgtacacgcattgggagcaaaagattaatgacgagaccgggaaattttggctccaggatcgactatgtcctttgcagtattgcaatga  
aggattggatctgacgacgccaacattcaggagggcctcatgggctctgatcattgccagtatatgccatgctcaaggatcgtgtcacg  
ctcaatggcgctcgagtatgatttgccagagctgatgaatcctcccggcacgtttcaagacggccaaaggaaaatcgaatggagcttcaa  
gagtttgcccgcgctctgtggacgactgcttccagagtttgaccgtcgaaggagcatcaaggacatgtttacgcgcaagaaaccggcct  
ctgctgcttctccgggtgcatctatctgttgcagcaactcattcagcaacttcttcaaagtgttcgcccgggtgttcagcatcccact  
gagcccgaaaccaaagtctgcatctgcaatatcgagtgaagccacagcaaccgtcagaacaaaaggcgccgttcgcgagcctactaaaag  
cccagagaagaaacgaaatgctgctggaggtcctaccgccaatccacaaaacggcagaagcttggccctgcttcgtctaaatcgaatc  
ccgcaaaggggtcaacagtcattgacgggattcctcaagccagccgcgagggcgacgaacaatggtagcgctgttaacgagcctgggttcg  
aggaagaccatgatagaaaaggaacgacaaaactcccgtcaatctctactccgaccctggaagcccgctctcccaagtatacgtcgtc  
tcagggcgctgctcgcaaatgacagaagacctcaatgggacgaccgcaaactctctaagaaccagtgaagagacggcagcctcagggg  
actaccgctctaccgcaatcgaatcagcctccaactcgtcagtggcactggaagatttcagcgatatcccagactcgcgagaaacgtgg  
tcgcggctctttgccaacgcgatccaccgcgctgcaagaccatggcgagttttgcatcagcctcacgaccaagaagagcggcgtcaa  
ctgtgggagagcattttggatatgtccgagaccaatcgggccccagcgggtccaaagaaaaagggacagaatggcgctgccccaccttca  
tctgggccagcgatttgaacggctccaccgcagcaccatga]

protein sequence =

[MPLRITSWNVNGIRNPFQYQPWRDKRSFDLLEADIVVMQELKIQRKDLTDDMVLVPGWDCYFSLPKHKKGYSGVAIYTRQSACVPIRA  
EEGVLGVLCPPGSSTSyrnLPADQAIGGYPTDEQIAAAGVDPAALDAEGRCVALEFPFVILGVYSPANSNGMRDDFRYGFLCALDARI  
RNLTALGKNIILTGDNLVSRQELDTANAEETLKKEGITHEEYVSSPNRRMFNQLLENGVEVGSRDEGREQPVLWDTCREFHPTKGMYT  
HWEQKINARPGNFGSRIDYVLCSIAMKDWICDANIQEGLMGSDHCPVYAMLKDRVTLNGVEYDLPELMNPPGTFQDGQRKIEWSFKSLP  
ALCGRLLPEFDRRRSIKDMFTRKKPASAASPGASISVAATHSATSSNGSPGVQHPTEPEPNAAASISSEPQQPSEQKAPFAEPTKSPEK  
KRNAAGGPTAKSTKRQKLGPASSKSNPAKGQQSLTGFLKPAARATNNGSAVNEPGSRKTMIEKERQTPVNLYSDPGSPSPKYTSSQGA  
ASQMTEDLNGTTANSLRTSEKTAASGDYRSTAIESASNSSVALEDFSDIPDSRETWSRLFAKRDPPrCEDHGEFCISLTTKSGVNCGR  
AFWICPRPIGPSGSKEKGTEWRCPTFIWASDWNGSTAAP]

start gene g2 : MAT1.1.4

```
unnamed-1AUGUSTUS gene 3403 5112 0.32 - . g2
unnamed-1AUGUSTUS transcript 3403 5112 0.15 - . g2.t1
unnamed-1AUGUSTUS stop_codon 3403 3405 . - 0 transcript_id "g2.t1"; gene_id
"g2";
unnamed-1AUGUSTUS terminal 3403 3717 1 - 0 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS internal 3814 3932 0.48 - 2 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS internal 4038 4631 0.95 - 2 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS initial 5067 5112 0.31 - 0 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS intron 3718 3813 1 - . transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS intron 3933 4037 0.48 - . transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS intron 4632 5066 0.58 - . transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS CDS 3406 3717 1 - 0 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS CDS 3814 3932 0.48 - 2 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS CDS 4038 4631 0.95 - 2 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS CDS 5067 5112 0.31 - 0 transcript_id "g2.t1"; gene_id "g2";
unnamed-1AUGUSTUS start_codon 5110 5112 . - 0 transcript_id "g2.t1"; gene_id
"g2";
```

coding sequence =

```
[atgttcgatatagacgacaggaccgccccttctcagccggctatcttccctccactctgatcctcggctccgacgcacagctggccgacctga
tccaagacttcgagacgctgcgaaaccaatacatcctcgacttcgggcaaccgacgagcagcgacctgcgcatctcgatcgagaagctg
gccgacgagtgccgcccgcgtgctgcgctccgcccgcctcatcgaggacgccccacgcgcccgcgcttctgcgcccgcgctccacaaact
gtcgaccgacctcatcgacttgacgaaagtcaagcgcaaggtccgcgacgtggtggccaagatcaaggaaggcgagcgcgctcaatagga
ataggaatgggtttgtcagggatttgccgcgctggagtcagtttttgataagaatgcgcccgggtgctagtagtccctgctgcttcagatg
cggcgcgctgctgcatcagaagcgcgatgcgtatcaactgattccgcggggcggcgttaagattcttcagttttgggacgaccggctcta
tgcgatgattcaggcaatcgacttggtgtggacatttctactcttctctcagctcccgcctaacatcgttgcagctgcggccagcacag
```

gcgtcgctttggcacaaggcaaacgagctaacgagcgaaacgcagcgcatcgagtggagccatccaagcatgcaccgcctcaactttggc  
ctgcgcacttttcgagcccatgacattcgagcccagcccatggaacctgctcgaccacgcgacccacaccttccgcgacgtatcgcgctt  
cgaccggcgctgggcgcgctgatgccgcgctcagcagccgcgacgagggccttccagccgcccgaccacgacgacgacgacgacgcct  
tcgacgccgtgctgacgcagcgcttcgtcaagcgctcctcctcagagaacaaggacgccgcggaccgcaaggaaaaggcggacaaggaa  
cgaaggcggaagatgcaggttgtgcagggggagggcccgcagcagcagcagcagggggcagcaacagggcgagagctcggcggtggcgctgc  
tgcgtag]

protein sequence =

[MFDRRQDRPSQPAISSTLILGSDAQLADLIQDFETLRNQYILDFGQPTSSDLRISIEKLADECRRVLR SARLIEDAPTRRRFCAALHK  
LSTD LIDLTKVKRKVRDVVAKIKEGERVNRNRNGFVRDLPRWSQFLDKNAPGASILLQMRVLHQKRDAYQLIPRGVKILQFWDDRL  
YAMIQAIDLVTFTLLSAPANIVAAAASTGVALAQGKRANERTQRIEWSHPSMHRNLNFGRLTFEPMTFEPSPWNLLDHATHTRDVS  
R FDRRLGALMPRVSSRDEAFQPPDHDDDDDAFDAVLTQRFVKRSSSENKDAADRKEKADKERRRKMVQVQGE GPQQQQGQQGESSAAGA  
AA]

start gene g3 : MAT1.1.1

```
unnamed-1AUGUSTUS gene 5353 7293 0.64 + . g3
unnamed-1AUGUSTUS transcript 5353 7293 0.64 + . g3.t1
unnamed-1AUGUSTUS start_codon 5353 5355 . + 0 transcript_id "g3.t1"; gene_id
"g3";
unnamed-1AUGUSTUS initial 5353 5704 0.64 + 0 transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS terminal 5840 7293 1 + 2 transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS intron 5705 5839 0.64 + . transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS CDS 5353 5704 0.64 + 0 transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS CDS 5840 7290 1 + 2 transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS stop_codon 7291 7293 . + 0 transcript_id "g3.t1"; gene_id
"g3";
```

coding sequence =

```
[atggctgtcgtgcaaaccttcctccccacgtcaagtcataacctggccgcatgctctcaagcccaaataatcagcttggtttcatcgc  
tcaaaggcgtcgaaggcaagaggacagtcgccgctgtggttgccgcgggcgaccctacaactcgtggctcgcactcgacgctctcttcc  
tccaagtcaacttccacccaagcgacaactccgtcagggtcacggtcgccttacatgactcccgcattcccgcacaaagaagacctcctc  
aagcgcgcaggcaaagaccctgccaaacttcgcagccagatagaatctctgcaaacatgaccaatgttcctcacgttgcgccagacagtt  
cgtccagtcgagctacatctcagacaggcgctacgcagacaagccctctacgagctctcaactctttcatggccttttcgagtttactac  
aatccgatcttccactcttttgagcagaaggagcgctcaggcttcatgaagataatgtggggccgaggaggaaaccaaagcgaagtggag  
catcattgccaaaggcatgggtcaatcatccgggatgaagtgggcaaagattctgctccgcttgacagattccttgagcttgccctgcctc  
acattgacatcctcgtcgcagatgcatacttgctcgaccttgggttgaggagcttgcgacggacaatggcgacacgtctcttagacggcaa  
ttcctgcccacttcaaggattttcctgccagcttccgcacaactacaaagtcagttcaggacgtcattgacttcgtccgcaacttggg  
atacgggtttacctaccaatgacacgcctccagcttcaagcaacatctcgacctctcatctgacgatggctgcacagcctgttcatcatt  
ccaccgcgaccacaggccctggattccttccctcttcgatattccactccacaaactccagctccaacgccaactttcagtctcttgca  
gcttccaatcttgagggcccccattgcgtccaacttccctgagctctgctatcagcagcacgacgcagacacacctcaatggcagcctcgac  
tacgctatcccacgctgggtgacaatgcttcagctccaattgaacaatctgcacaggaaaatgcaatggatcttggcagtgacgcgcagc  
aatcagcggcaggcaccaatcctagttatgcggtcagccagcagcaggcaagcaatgcaacaaacgcaccacctcctcaaacaagcagc  
gtcaatgcagcttctgggcttacgtccggtagcaatgtggcagcgcgaatccaatgaccaggctctgatggccatcatccaccaacatgc  
tgctgctgccaatgctggtaatggaggcctgctcactggaggcccatgactcaccagcagatgttgacttacatgtatggggccaacg  
tacaaccagcccaccccgacatgtactcaccacacgtggacgatacgattgcgaatcccgacccccagagttgctgtatgcagcgcagcc  
agcctcgccgatccagatcttcttgccatctccatcaacgcggaataccctcatgaaatcctgaactattaccctgcttcaacgcgcaa  
ttcaggagaatcgatggcaagctcgggttgctgcgcccgcacctgcgcagcaggacacgcagaccgcctggctggactgcgccctacagatca  
tggaacacatcaccatcatcttcaccaacatttcccttatgccgagttcttcaatcctgacctgagggcttgaaatgtcttcgacattgac  
ggaattcccgatctttactgccgttgacaacaccgacttccagttccaggagatgtccttgcccttgggtcaccagccactgggatgatga  
gtccgatgggtccagagcagtgatga]
```

protein sequence =

```
[MAVAANLP PHVKS YLAACSQAQINQLVSSLKGVEGKRTVA AVVAAGDP TTRGSHSTLSSSKSTSTQATTPSGSRSPYMTPAFPTKKTS  
SSAQAKTLPTSQPDRISANMTNVP HVAPDSSSSRATSQTGATQTSPLRALNSFMAFRVYYNPIFHSFEQKERSGFMKIMWAE EETKAKW
```

SIIAKAWSIIRDEVGKDSAPLDRFLELACPHIDILARDAYLSTLGWELATDNGDTSLRRQFLPNFKDFPASFRRTTTSVQDVIDFVRNL  
GYGLPTNDTPPASSNISTSHLTMAAQPVHHSTATTTGPGFLPSSIFHSTNSSSNANFQSLAASNLGPHASNFPESAISSTTQTTSMAAS  
TTLSHAGDNASAPIEQSAQENAMDLGSDAQQSAAGTNPSYAVSQQQASNATNAPPPQTSSVNAASGLTSGSNVAAQSNDQALMAIIHQH  
AAAANAGNGGLLTGGPMTHQQMLTYMYGANVQPAHPDMYSPHVDDTIANPDPQSCVMQRASLADPDLLAISINAEYPHEILNYYPASTR  
NSGESMASSVVRPTCDEDTQTAWLDCALQIMDTSPSSSPTFPYAEFFNPDHEGLNVFDIDGIPDFTAVDNTDFQFQEMSLPLVTSHWDD  
ESDGPEQ]

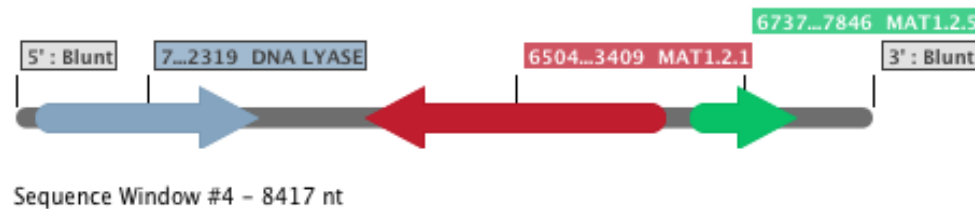


FIGURA 2:  
MAT1.2.5 DE *P. CITRICARPA*

DNA LYASE + MAT1.2.1 +

>*P. citricarpa* DNA LYASE + MAT1.2.1 + MAT1.2.5

```

ATGCCGCTCCGGATAACCTCGTGGAACGTCAATGGCATCCGCAATCCTTTTGGCTACCAGCCATGGCGCGACAAGCGGTCCTTCGACGC
CATGTTTCGACCTCCTCGAGGCCGACATTGTTCGTTCATGCAAGAGCTCAAGATTCAGCGCAAGGACCTCACCGACGACATGGTCCTGGTTC
CCGGCTGGGATTGCTACTTTAGTCTGCCAAAGCACAAAAAGGTGCGCGTGCTGCCGCTCTGATGTGACGGCTGCTGACGGCTTCAGGA
TACTCTGGCGTAGCGATATACACTCGCCAATCTGCCTGCGTGCCCATTCGCGCTGAAGAGGGCGTCCTGGGCGTCCTCTGCCCTCCAGG
ATCCTCCACGTCCTACCGCAATCTGCCCCGCCGACCAAGCCATTGGCGGATATCCAACCGACGAACAAATTGCCGCCGCCGGCGTCGACC
CTGCTGCGTTGGATGCCGAAGGGCGCTGCGTCGCTCTCGAGTTTCCTGCCTTTGTCATTCTCGGCGTCTACAGCCCGGCGAACTCGAAT
GGAATGAGAGACGACTTCCGCTACGGGTTCTCTGCGCCTTGGACGCGCGCATACGTAATTTGACGGCGTTGGGCAAGAATATCATTCT
GACTGGCGACCTCAATGTCTCGCGGCAAGAACTGGACACGGCGAATGCGGAGGAGACGCTCAAGAAGGAGGGCATCACGCACGAAGAGT
ACGTGTCGTCGCCAACAGGCGCATGTTTAACCAGCTGCTGGAAAATGGCGAGGTTGTGGGTAGTCGAGACGAAGGACGGGAGCAGCCG
GTGCTGTGGGACACGTGCCGAGAGTTTCATCCTACGCGCAAGGGCATGTACACGCATTGGGAGCAAAAGATTAATGCGAGACCGGGAAA
TTTTGGCTCCAGGATCGACTATGTCCTTTGCAGTATTGCAATGAAGGATTGGATCTGCGACGCCAACATTAGGAGGGCCTCATGGTTA
GTTGATGCGCTTCGATGTCTTGAACGACTGCTGACTTTTGTTCAGGGCTCTGATCATTGCCAGTATATGCCATGCTCAAGGATCGTGT
CACGCTCAATGGCGTCGAGTATGATTTGCCAGAGCTGATGAATCCTCCCGGCACGTTTCAAGACGGCCAAAGGAAAATCGAATGGAGCT

```

TCAAGAGTTTGCCCGCGCTCTGTGGACGACTGCTTCCAGAGTTTGACCGTCGAAGGAGCATCAAGGACATGTTTACGCGCAAGAAACCG  
GCCTCTGCTGCTTCTCCGGGTGCATCTATTTCTGTTGCAGCAACTCATTAGCAACTTCTTCAAATGGTTTCGCCCCGGTGTTCAGCATCC  
CACTGAGCCCGAACCAAATGCTGCATCTGCAATATCGAGTGAGCCACAGCAACCGTCAGAACAAAAGGCGCCGTTTCGCGGAGCCTACTA  
AAAGCCCAGAGAAGAAACGAAATGCTGCTGGAGGTCCTACCGCCAAATCCACAAAACGGCAGAAGCTTGGCCCTGCTTCGTCTAAATCG  
AATCCCGCAAAGGGTCAACAGTCATTGACGGGATTCCTCAAGCCAGCCGCGAGGGCGACGAACAATGGTAGCGCTGTAAACGAGCCTGG  
TTCGAGGAAGACCATGATAGAAAAGGAACGACAAACAACCTCCCGTCAATCTCTACTCCGACCCTGGAAGCCCGTCTCCCAAGTATACGT  
CGTCTCAGGGCGCTGCGTCGCAAATGACAGAAGACCTCAATGGGACGACCGCAAACCTCTCTAAGAACCAGTGAAAAGACGGCAGCCTCA  
GGGGACTACCGCTCTACCGCAATCGAATCAGCCTCCAACTCGTCAGTGGCACTGGAAGATTTAGCGATATCCCAGACTCGCGAGAAAC  
GTGGTCGCGGCTCTTTGCCAAACGCGATCCACCGCGCTGCGAAGACCATGGCGAGTTTTTGCATCAGCCTCACGACCAAGAAGAGCGGCG  
TCAACTGTGGGAGAGCATTTTTGGATATGTCCGAGGTAAGCACTTCGCCCAGTTGTGGCAAGCATGATTAGTTTGTTCGCATTGTCAATGT  
CATACTGGCTTTTACCATCTTGGTGCCACGCATTTACGATTACATCTTTCCAATGTGCAGCTTCTTGATTGGAATGCTGCATGACCTTT  
CCCCTCCCGCCATTCACCTCTTTCCAAGACGCAATCAAGTTGAGGCTTTTTGTATGCTGACCCTCCTTGCTTTTTTGCAGACCAATCGGCCC  
CAGCGGGTCCAAAGAAAAAGGGACAGAATGGCGCTGCCCCACCTTCATCTGGGCCAGCGATTGGAACGGCTCCACCGCAGCACCATGAA  
GGCGAGTGTCCAAAGGCACGACGGCAGATTTATGTGTGGAAGAGCGCAGCAGAGAATGATTTCTCGCGCATATCGGACCAACCCACCGG  
GTGACCCATCACATCGACGCCTCCGTAGCTGAAGTCTGCCGATCCCTGAAGGTGCCCCACTACAGCCCTGTCGTGGCAGGTTCTTCCCG  
AACTGTCTTGCCCTTGCTAGACCGTCAGACAACTTGCCTACGATTTATGAAACATGCCCTCTCCATGGTCCGGTTGTCCATGCGTTTCT  
CTGTCTGCAAACCTGCGGGTAGGATCCCGTGTATCAACAAGTGTACAATTGACGGGGCCGAGCTCGGGGTGTGTTTGTTCCTCTGTTTC  
CCATGTCTCGTTTTTCCGTCTTTCAGAGAGGTGCGACAGGCAACAGGGGAAGAGAGACAGGCGCATGAATCAAGACAACGGCGCCACGAG  
ACAGGGCAAGACAGAAGAACTGGTGGCACACGTGGCTACGAGACAAGCCGTGTTGCTGGCACGGCTACGGCATGGCAGACCATGGCGA  
AATGGGTTTCCGTCCACAACGCGCCCTGCCACGGCGTGCATCTCTGCCGCTCGTCCAGTGCACCGGCCACCAGCTAGCTGATCGACTTA  
CACAGTGACGTATCCAGCGTTTGCATTGCGCAGATTGGATTCTGAGAGTTACGACTGCCGCCTTCACCGCCTTCACCGCCGCCATACCA  
AACATTGATGTCGCGATGTTTCGTAAATTCGTAGACTGCAAAACAAAATCGCTACGTATGTAGCTCTTCCCCATGCGCCCCAACGCCGTC  
GGTGCAGCCGTCATATGCCAGCCCTGCCACGCGCTTGCTCGTAGCCCTCCCGCCGCTCGCGGGCGCCGCTTGCGTGGAAGAAAGCGAATG  
CAAGAGAAGAGATGGCTAGTTTGTGTTTTTGGCCGCCTTCTCACGTCGTCCTTCTCCTCCGCTGCTCGCGCCCAGCAGCATCCGCGTCC  
CGATATCCCTCCTTCAAACGAAGCCATCCAACCTATCTCCGCGTCGACCTCGCAACGTACATAGCCATCCATCCCACACACCATCACCA  
CGCCACACCTCCACCTCCATCTACGCAGCAGCGCCAGCCGCCGAGCTCTCGCCCTGTTGCTGCCCTGCTGCTGCTGCTGCGGGCCCTCCCC

TGTGCACAACCTGCATCTTCCGCCTTCGTTCCCTTGTCCGCCTTTTTCCTTGCGGTCGCGGGCGTCTTGTCTCTGAGGAGGAGCGCTTGA  
 CGAAGCGCTGCGTCAGCACGGCGTCGAAGGCGTCGTCGTCGTCGTCGTCGGTCGGGCGGCTGGAAGGCCTCGTCGCGGGCTGCTGACGCGC  
 GGCATCAGCGCGCCCAGGCGCCGGTTCGAAGCGCGATACGTCGCGGAAGGTGTGGGTGCGGTGGTCGAGCTGTGAGTGGGAGAGGGGGGT  
 GTTAGTGCGGTGTTTGTGTGGGTGGTGGGTGGGAGGGTTGGGGATGGGAGTATGGCGTCTGTGGACGTACGTACCAGGTTCCATGGGC  
 TGGGCTCGAATGTCATGGGCTCGAAAGTGCGCAGGCCAAAGTTGAGGCGGTGCATGCTTGGATGGCTCCACTCGATGCGCTGCGTTCGC  
 TCGTTAGCTCGTTTGCCTCTTCGTCGTCATCATCGGCATCAGCATCTGCTGCCTGTGTGCTCCTCCCCTTCTCTTTCTCTTCTGCTAT  
 GCCGAGACGCGGTCTAGAGGAGCCAAAACGTACCTTGTGCCAAAGCGACGCCTGTGCTGGCCGCAGCTGCAACGATGTTAGCGGGAGCTG  
 AGAGAAGAGTAGGAAATGTCCACACCAAGTCGATTGCCTGAATCATCGCATAGAGCCGGTCGTCCCAAACTGAAGAATCTTAACGCCG  
 CCCCgcggaATCAGTTGATACGCATCGCGCTTCTGATGCAGCACGCGCCGCATCTGAAGCAGCAGGATACTAGCACCCGGCGCATTCCTT  
 ATCCAAAACTGACTCCAGCGCGGCAAATCCCTGACAAACCCATTCCCTATTCCCTATTGACGCGCTCGCCTTCCTTGATCTTGGCCACCA  
 CGTCGCGGACCTTGCGCTTGACTTTCGTCAAGTCGATGAGGTGCGTCGACAGTTTGTGGAGCGCGGGCGCAGAAGCGGCGGGCGCGTGGGG  
 GCGTCCTCGATGAGGCGGGCGGAGCGCAGCACGCGGCGGCACTCGTCGGCCAGCTTCTCGATCGAGATGCGCAGGTGCTGCTCGTCGG  
 TTGCCCgaAGTCGAGGATGTATTGTTTTCGCAGCGTCTCGAAGTCTTGGATCAGGTGCGCCAGCTGTGCGTCGGAGCCGAGGATCAGAG  
 TGGCTGCGGTTTCGTACGCGGGTTGATGTGGGAATTGGCGCGTGTGCTTGCCTTGTGAGCCTGTGAATCCGGAGGCCTTGAGGCCTGCT  
 TCTGGATCTTCGATTCCGGAGGCGCGGTTTTTTTCTGGCGCGCTCGAGGAGAGGCGCTTGTGGGTTCAGAGGTCCATGGCTTGTGGGG  
 TGTGATGGGCAAAGGTTGATCTCTAGGTGCCTGGCTCTATCGGGGGAAGAAAGGAATCAGCGCTTTCGGAGCGTCGTCTTCGTTGGTCA  
 AAATTTcggggGAGCTAGTGCGGGCTTTGAGATTGGATTTCTATCTACTACTATGACTAACCgCGCCATCAGTCCAAACTTTTCATGCT  
 CTCTTTTGTTTTTTGTCTTGTGTTTCCACGGGCAACAAATGACGCTGGTGCTCAAATGTGCTTGATGTCTCCAAAGTCCAGTTCAGTGC  
 CGCCCAAGAAATCATAGTTTAGGTGCGCAAGTGGAATTTGAAGCCTTCGTTTCATAGGTTcAGCACTGAAGCTGACCTCACGGTCAGCC  
 GCCGCATTGAAGCTGTGCTCAACGAAGGACTGGTCAGTGGCAGACTGAACACTGAAGCCAAAATTGTCCAGAAGAGCATGTCTGAGCTG  
 TTTCTCTGCGACGTCCATGGCTTCAGCCTGCTGTTGCAGGCCTACAGAATCGAGAGTTGCTTCGTAAGCGTCTTGATCGGTAGCCGCGT  
 TGGAGGACAGGGTGACAAGAGGGTGTGTTGGTTGGCGCGGTTGGAAATCCCGTATGGTGGCTGAGAGTTGTCTGAAACATCTGCCTGCC  
 GTCGCGCTTGGATCCAAGTCATATCCTTGGTATTTCTGGTGTGGGTGACGAGATATGCGGCTGTCTGTCCAGAAGAAATCGGAAGAGA  
 GGTGGGAGGGTCGAACGTGAGAGGCTCGTCAGAAATCGAGGGCCTGTTGGCCGCAGCTGCAGCAGCTTCAGCCAGCTTCTTCTTGGTCA  
 TCCGTCTCTTCTTCTCGCCAGGCTTGCCTGGACGATACTGGTAGTCGGGAAATCTGAGGTGGTGCTCTTCTTTGCCAGCTTGGCTCGC  
 TCACGCCAAAACTCCTTCGCTTCGGGCGACTCTTCTTGCCACATCTTGCCTATGATTTTGGCTGATAACATCGTcAGTGTTAATTGTTG



CGCCGAGCAGGTGAAGTTCTTACAAATTTGATTGTTGTGCAAGCCTGGATTTTCCATCGTGACCCTCTGGTGAAGAAAGTTGCGATACA  
TAATGAAAGGATTGCAAGGACGCGGCACCTTGGGCTTGGAAAGGTCGCGATGTAGCATCCAGGACTCGAGTCCGAGTGCTCTTCGCTTCG  
ATCGTGGCAGGAGGCTTTGCAGGAGTCGCATTCTTGTCGACCGCGGACGGTGCGTTACCTCAGACGGCTGAACAGTCCTGAGCTCGGG  
GGGCGGAAGAATACGCAAGACCTTCTGCTCTGCGTCCACAGTCTTCTCGACCGGAGCCTTGACATGTTGCTAACATTGGTCAGCAAGT  
TAGATCACTGAGCGACAGAGGGTTGGTACTGTACCTGAAGCGTTGAGCGACACGATCAACTCCGGCTTGTCCAAGAGCGTACGCAATTT  
CGAATGGGACGACGACTTCGATGCGACCTTTGCGAATTTGCTTGAGCGACTCAGCCCAGAAACGATCCACCAGCTGGTTTCACGGCGCGC  
ATCACGAAGGCAGGTCCACCGGCCGGGACTGGAACGTGTGGAGAGCTCATTTTGGTCTGAGTTGGTTGCAGGAGGTTGAAGTCGAGGGG  
CTGTTGTGTGCGAAGTTGTGTTGAGGTCTGAGGGTCTGAGGGTTGGGGTTGTTGAAAGGGTGGGGAAGGGTCTTGAGTAGAGGAAAACA  
TTCTGGCGCTGGCGCAGGCGCAACAAAACCTCTATCAAGCGACCCTCAGTTCACCTCACACGGACCTCTAACCTACACATTGACTTGTCA  
GCGAAACAGAGTGAATGGCATTACAAACCTGAGGCTTGTGCATTTGCACTTACCGATCGCTGGTCGTTCCACACAGCTCCTGCTCAGC  
TCGACACTCCAACATCTTGCCTTCAAATTCTCCACCGATTCTCTTCTCACAGCTATGGACAACAACGAAATTATCCGCCTCCTCGCTC  
TGCAAGATGATTACCAACCAACTGCTGAGGCTCCCGAGGAGAATTCCTTTTGTGCTGCCGAAAAGCCCCAGGATTTCTTCCCACCAACA  
GTCGATGATCCGGATGACAACGGTACTTCGCCAGCAGGGCGTATCTTCGAGCCCTTCAATAAGAACCGTCTTGAATGGGAGGCGTTCTGA  
CCGATCAAAATCTTGCGAAAACGACCAAGCAATAGCCGACAAGCTCAAAGAGCTGAATTCAGATGTTGAAGAATCGAGTGGCTTCAACT  
TCAACAATTCTATGGTTGATAGCCCGGAAAGACAATGTGCGATGAACCGAACCTTTCTCTTCCAAGAAGGCGTCGATCCACGTTTTTCAG  
AGACTTGACGACTTGACGATGCTTTACCGCATGGAAGAGGGACTGGCAAACAACATGTTAACAATGATCAATGATGGCGCTTCAAAGGA  
ATTCTCAAGATGCATAGGTCAATCAGAGGAAGAGCCCCGAGGGGCAGAGATTTTCGCCGCGAGGCAATCAATCACTGTGCCTTGAAGCAAG  
CACTTACTGAGTAAGTCGTGTATGAAGACACAAGCGACAGTATACTGATATTCCACAGAACTACCGATGCGACCGCGCGAAGCTGCAAT  
CTCGTCAAGCTCGTCGTCGAGCCGTATGAGGTGGAGGCCGATTTTCGCGAACCTGGACTTTCCAGAGGAACTTGGCGACAATCTTTGGCA  
GGAGGATGTGACTCGCTCGAGATGGACTTCGCGAAGGGACGTCGAAGCAAACGATACGCTTGAAAGCTTCTTCAGCATATGCCAGGAGC  
CCAACGTGCTTGAAGTCGATTTCTTGTCTCAGTACCTGGGATTGGAGACAGAGAGCGTGACGACTGGTGTAAGTGGCTTACCAACTAC  
ATTCAGCGCATGGCTTAAAAGAAAGAGTGGCTAACCTTCCTTTTCAGTCAGGGGCAAGCAGAAGAGGGTCAAGCATCTCTTTGTTACAAG  
AAAGCTGCCCCAGGGCAACAGCGTTCGCATGTTGGACCACCGCATGCGGCAAAGTCTGCACGACATTTTCCCTCCGCGTCTGAGCATTT  
CTAGATAAGCTTTACAGAAATGAGGTGATTTCATAGCGATACTGGGGAAGAAAAGTTTGGATGCGGTGGCGGAGAAGCGTGGCTGATGT  
CTCAGTCTTGGGCATCAGGCACAGACCAGAGTGGCGTTTTTCCACGGAGCAAAATCGTGTTTCTTTTCATCTTCAAGTCAGCATCGCTGA  
GGTGGGCACGCGGCGCATCGTTCGTGGGCTCATATGGATGGTACATCTTCTCATTTGCAAGAGTAGTTAACTTTGACAAGCGCGTGAGG

TGGTGTGATCCGGAAAACTTCAGCGCCTAGCGCCCAGTATTGCCCAGACCCGGCCCCGCCCGGTTGCTACCCGCGTGCGGCGCAAGTA  
CGATAGTGCATTGACACTAGGGACCCCACACCCGCATCCGCCCCAACACATCAAAGCGACCTATACACCACCAGCATCTCGAAAACTTGC  
CGTTCCATCGCGCCGCCACAGAATCTTCTCTGCCTGGCTGTTTTTTGGAGACTGCGCCGAACAGCTCAGGCACGCCACAGGAGCCAAG  
CACACGTGGCTCAGAAGGCCCTCTCAGTCTGCGGCTTACAACCCCCCCCCC

Predicted genes for sequence number 1 on both strands

start gene g1 : DNA LYASE

unnamed-1AUGUSTUS	gene 7	2319	0.56	+	.	g1
unnamed-1AUGUSTUS	transcript	7	2319	0.27	+	g1.t1
unnamed-1AUGUSTUS	start_codon	7	9	.	+	0 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	initial	7	226	0.81	+	0 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	internal	272	981	0.55	+	2 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	internal	1031	1998	0.7	+	0 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	terminal	2220	2319	0.74	+	1 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	intron	227	271	0.55	+	. transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	intron	982	1030	0.91	+	. transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS	intron	1999	2219	0.74	+	. transcript_id "g1.t1"; gene_id "g1";

```

unnamed-1AUGUSTUS CDS 7 226 0.81 + 0 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 272 981 0.55 + 2 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 1031 1998 0.7 + 0 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS CDS 2220 2316 0.74 + 1 transcript_id "g1.t1"; gene_id "g1";
unnamed-1AUGUSTUS stop_codon 2317 2319 . + 0 transcript_id "g1.t1"; gene_id
"g1";

```

coding sequence =

```

[atgccgctccggataacctcgtggaacgtcaatggcatccgcaatccttttggctaccagccatggcgcgacaagcggtccttcgacg
ccatgttcgacctcctcgaggccgacattgtcgtcatgcaagagctcaagattcagcgcgaaggacctcaccgacgacatggtcctgggt
cccggctgggattgctacttttagtctgccaaagcacaaaaaaggatactctggcgtagcgatatacactcgccaatctgcctgcggtgcc
cattcgcgctgaagagggcgctcctgggcgtcctctgccctccaggatcctccacgtcctaccgcaatctgcccgcgcgaccaagccattg
gcggatatccaaccgacgaacaaattgccgcccgcggcgctcgaccctgctgcgttggtatgccgaagggcgctgcgtcgctctcgagttt
cctgcctttgtcattctcggcgtctacagcccggcggaactcgaatggaatgagagacgacttccgctacggggttcctctgcgccttgga
cgcgcgcatacgtaatttgacggcggttgggcaagaatatcattctgactggcgacctcaatgtctcgcggaagaactggacacggcgga
atgcggaggagacgctcaagaaggagggcatcacgcacgaagagtacgtgtcgtcgcccaacaggcgcatgtttaaccagctgctggaa
aatggcgaggttgtgggtagtcgagacgaaggacgggagcagccggtgctgtgggacacgtgccgagagtttcattcctacgcgcaaggg
catgtacacgcattgggagcaaaagattaatgcgagaccgggaaattttggctccaggatcgactatgtcctttgcagtattgcaatga
aggattggatctgcgacgccaacattcaggagggcctcatgggctctgatcattgcccagtatatgccatgctcaaggatcgtgtcacg
ctcaatggcgctcgagtatgattttgccagagctgatgaatcctcccggcacgtttcaagacggccaaaggaaaatcgaatggagcttcaa
gagtttgcccgcgctctgtggacgactgcttcagagtttgaccgtcgaaggagcatcaaggacatgtttacgcgcaagaaaccggcct
ctgctgcttctccgggtgcatctatcttctgttgacgcaactcattcagcaacttcttcaaatagggtcgcccgggtgttcagcatcccact
gagcccgaaccaaataatgctgcatctgcaatatcgagtgcacacagcaaccgtcagaacaaaaggcgccggttcgcggagcctactaaaag
cccagagaagaaacgaaatgctgctggaggtcctaccgccaataccacaaaacggcagaagcttgccctgcttcgtctaaatcgaatc
ccgcaaaggggtcaacagtcattgacgggattcctcaagccagccgcgagggcgacgaacaatggtagcgctgttaacgagcctgggttcg

```

```

aggaagaccatgatagaaaaggaacgacaaacaactcccgtcaatctctactccgaccctggaagcccgtctcccaagtatacgtcgtc
tcagggcgctgcgtcgcaaatacagaagacctcaatgggacgaccgcaaactctctaagaaccagtgaaaagacggcagcctcagggg
actaccgctctaccgcaatcgaatcagcctccaactcgtcagtgggactggaagatttcagcgatatcccagactcgcgagaaacgtgg
tcgcggtctctttgccaaacgcgatccaccgcgctgcaagaccatggcgagttttgcatcagcctcacgaccaagaagagcggcgctcaa
ctgtgggagagcattttggatatgtccgagaccaatcgggccccagcgggtccaaagaaaaagggacagaatggcgctgccccaccttca
tctgggccagcgattggaacggctccaccgcagcaccatga]

```

protein sequence =

>DNA LYASE MAT

```

MPLRITSWNVNGIRNPFQYQWRDKRSFDAMFDLLEADIVVMQELKIQRKDLTDDMVLVPGWDCYFSLPKHKKGYSGVAIYTRQSACVP
IRAEGLVGLVLCPPGSSTSYRNLPADQAIGGYPTDEQIAAAGVDPAALDAEGRCVALEFPFVILGVYSPANSNGMRDDFRYGFCLCALD
ARIRNLTALGKNIILTGDNLNVSQELDTANAEETLKKEGITHHEYVSSPNRRMFNQLLENGVEVGSRDEGREQPVLWDTCREFHPTKRG
MYTHWEQKINARPGNFGSRIDYVLCIAMKDWICDANIQEGLMGSDHCPVYAMLKDRVTLNGVEYDLPELMNPPGTFQDGQRKIEWSFK
SLPALCGRLLPEFDRRRSIKDMFTRKKPASAAASPGASISVAATHSATSSNGSPGVQHPTEPEPNAASAISSPEQQPSEQKAPFAEPTKS
PEKKRNAAGGPTAKSTKRQKLGPASSKSNPAKGQQSLTGFLKPAARATNNGSAVNEPGSRKTMIEKERQTPVNLYSDPGSPSPKYTSS
QGAASQMTEDLNGTTANSLRTSEKTAASGDYRSTAIESASNSSVALEDFSDIPDSRETWSRLFAKRDPPRCEDHGEFCISLTTKKSGVN
CGRAFWICPRPIGPSGSKEKGTEWRCPTFIWASDWNGSTAAP]

```

```

unnamed-1AUGUSTUS transcript      3409 6504 0.09 -      .      g2.t2 : MAT1.2.1
unnamed-1AUGUSTUS stop_codon      3409 3411 .      -      0      transcript_id "g2.t2"; gene_id
"g2";
unnamed-1AUGUSTUS terminal 3409 3723 1      -      0      transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS internal 3820 4001 0.36 -      2      transcript_id "g2.t2"; gene_id "g2";

```

unnamed-1AUGUSTUS	internal	4044	4637	0.93	-	2	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	internal	5161	5751	0.79	-	2	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	internal	5815	6128	0.63	-	1	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	initial	6182	6504	0.45	-	0	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	intron	3724	3819	1	-	.	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	intron	4002	4043	0.36	-	.	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	intron	4638	5160	0.99	-	.	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	intron	5752	5814	0.6	-	.	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	intron	6129	6181	0.93	-	.	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	CDS	3412	3723	1	-	0	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	CDS	3820	4001	0.36	-	2	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	CDS	4044	4637	0.93	-	2	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	CDS	5161	5751	0.79	-	2	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	CDS	5815	6128	0.63	-	1	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	CDS	6182	6504	0.45	-	0	transcript_id "g2.t2"; gene_id "g2";
unnamed-1AUGUSTUS	start_codon	6502	6504	.	-	0	transcript_id "g2.t2"; gene_id "g2";

coding sequence =

[atgttttctctactcaagacccttccccaccctttcaacaaccccccaaccctcagaccctcagacctcaacacaacttcgacacaac  
agcccctcgacttcaacctcctgcaaccaactcagaccaaaatgagctctccacacgttccaggtcccggccggtggacctgccttcgtg  
atgcgcgcgctgaaccagctgggtggatcggtttctgggctgagtcgctcaagcaaattcgcaaaggtcgcatcgaagtcgtcgctccatt  
cgaaattgcgtacgctcttggacaagccggagttgatcggtgctcgctcaacgcttcagcgaacatgtcaaggctccggtcgagaagactg  
tggacgcagagcagaaggtcttgctgatttcttccgccccccgagctcaggactgttcagccgtctgaggtgaacgcaccgtccgcggtc  
gacaagaatgcgactcctgcaaagcctcctgccacgatcgaagcgaagagcactcggactcgagtcctggatgctacatcgcgaccttc  
caagcccaaggtgccgcgtccttgcaatcctttcattatgtatcgcaactttcttcaccagaggggtcacgatggaaaatccaggcttgc  
acaacaatcaaatttgcaagatgtggcaagaagagtcgcccgaagcgaaggagttttggcgtgagcgagccaagctgggcaaagaagag  
caccacctcagatttcccgactaccagtatcggtccacgcaagcctggcgagaagaagagacggatgaccaagaagaagctggctgaagc  
tgctgcagctgcggccaacaggccctcgatttctgacgagcctctcacgttcgaccctcccacctctcttccgatttcttctggacaga  
cagccgcatactctgctcgacccacaccagaaataccaaggatatgacttggatccaagcgcgacgggacaggcagatgtttcagacaact  
ctcagccaccatacgggatttccaaccgcgccaaccaacaacccctcttgtcacctgtcctccaacgcgggtaccgatcaagacgctta  
cgaagcaactctcgattctgtaggcctgcaacagcaggctgaagccatggacgtcgagagaaacagctcagacatgctcttctggaca  
atthttggcttcagtgttcagtctgccactgaccagtccttcggtgagcacagcttcaatgcggcggtgacccactctgatcctcggc  
tccgacgcacagctggccgacctgatccaagacttcgagacgctgcgaaaccaatacatcctcgacttcgggcaaccgacgagcagcga  
cctgcgcacatctcgatcgagaagctggccgacgagtgccgcccgcgtgctgcgctccgcccgcctcatcgaggacgccccacgcgcccgc  
gcttctgcgcccgcgtccacaaaactgtcgaccgacctcatcgacttgacgaaagtcaagcgcaaggtccgcgacgtggtggccaagatc  
aaggaaggcgagcgcgctcaataggaataggaatgggtttgtcagggatttgccgcgctggagtcagtttttgataagaatgcgcccggg  
tgctagtatcctgctgcttcagatgcggcgcgctgctgcatcagaagcgcgatgcgtatcaactgattccgcggggcgcggttaagattc  
ttcagttttgggacgaccggctctatgcatgattcaggcaatcgacttgggtgtggacatttcttctctcagctcccgcctaac  
atcgttgcagctgcggccagcacaggcgctcgctttggcacaagagaaagagaaggggaggagcacacaggcagcagatgctgatgccga  
tgatggacgacgaagaggcaaacgagctaacgagcgaacgcagcgcatcgagtggagccatccaagcatgcaccgcctcaactttggcc  
tgcgcaactttcgagcccatgacattcgagcccagcccatggaacctgctcgaccacgcgacccacaccttcgcgacgtatcgcgcttc  
gaccggcgcttggcgcgctgatgccgcgctcagcagccgcgacgaggccttccagccgcccgaccacgacgacgacgacgacgcctt  
cgacgccgtgctgacgcagcgcttcgtcaagcgctcctcctcagagaacaaggacgccgcggaccgcaaggaaaaggcggaagaaggaac  
gaaggcggaagatgcagggtgtgagggggaggggcccgacgacgacgagggggcagcaacaggggcgagagctcgggcggtggcgctgct

gcgtag]

protein sequence =

[MFSSTQDPSPPFQQPPTLRPSDLNTTSTQQPLDFNLLQPTQTKMSSPHVPVPAGGPAFVMRAVNQLVDRFWAESLKQIRKGRIEVVVP  
FEIAYALGQAGVDRVAQRFSEHVKAPVEKTVDAEQKVLRLPPPELRTVQPSEVNAPSAVDKNATPAKPPATIEAKSTRTRVLDATSRP  
SKPKVPRPCNPFIMYRNFLHQRVTMENPGLHNNQICKMWQEESEPEAKEFWRERAKLGKEEHHLRFPDYQYRPRKPGEKKRRMTKKKLAE  
AAAAAANRPSISDEPLTFDPPTSLPISSGQTAAYLVDPHQKYQGYDLDP SATGRQMFQTTLSSHHTGFPTAPTNNPLVTLSSNAATDQDA  
YEATLDSVGLQQQAEAMDVAEKQLRHALLDNFGFSVQSATDQSFVEHSFNAAADPTLILGSDAQLADLIQDFETLRNQYILDFGQPTSS  
DLRISIEKLADECRRVLR SARLIEDAPTRRRFCAALHKLSTD LIDLTKVKRKVRDVVAKIKEGERVNRNRNGFVRDLPRWSQFLDKNAP  
GASILLQMRVLHQKRDAYQLIPRGGVKILQFWDDRLYAMIQAIDL VWTFTP TLLSAPANIVAAA STGVALAQEKEKGRSTQAADADA  
DDGRRRGKRANERTQRIEWSHPSMHLNFGRLTFEPMTFEPSPWNLLDHATH TFRDVS RFRRLGALMPRVSSRDEAFQPPDHDDDDDA  
FDAVLTQRFVKRSSSENKDAADRKEKADKERRRKMQVVQGE GPQQQQGQQQGESSAAGAAA]

end gene g2

###

start gene g3 : MAT1.2.5

unnamed-1AUGUSTUS gene 6737 7846 0.81 + . g3

unnamed-1AUGUSTUS transcript 6737 7846 0.81 + . g3.t1

unnamed-1AUGUSTUS start\_codon 6737 6739 . + 0 transcript\_id "g3.t1"; gene\_id  
"g3";

unnamed-1AUGUSTUS initial 6737 7314 0.88 + 0 transcript\_id "g3.t1"; gene\_id "g3";

unnamed-1AUGUSTUS internal 7363 7640 0.92 + 1 transcript\_id "g3.t1"; gene\_id "g3";

unnamed-1AUGUSTUS terminal 7707 7846 0.92 + 2 transcript\_id "g3.t1"; gene\_id "g3";

```

unnamed-1AUGUSTUS intron    7315 7362 0.87 +    .    transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS intron    7641 7706 0.93 +    .    transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS CDS      6737 7314 0.88 +    0    transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS CDS      7363 7640 0.92 +    1    transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS CDS      7707 7843 0.92 +    2    transcript_id "g3.t1"; gene_id "g3";
unnamed-1AUGUSTUS stop_codon 7844 7846 .    +    0    transcript_id "g3.t1"; gene_id
"g3";

```

coding sequence =

```

[atggacaacaacgaaattatccgcctcctcgctctgcaagatgattaccaaccaactgctgaggctcccgaggagaattccttttgtg
ctgccgaaaagccccaggatttcttcccaccaacagtcgatgatccggatgacaacgggtacttcgccagcagggcgatatcttcgagccc
ttcaataagaaccgtcttgaatgggaggcgttcgaccgatcaaaatcttgcgaaaacgaccaagcaatagccgacaagctcaaagagct
gaattcagatgttgaagaatcgagtggttcaacttcaacaattctatggttgatagcccgggaaagacaatgtcgatgaaccgaacct
ttctcttccaagaaggcgtcgatccacgttttcagagacttgacgacttgacgatgctttaccgcatggaagaggggactggcaaacaac
atgttaacaatgatcaatgatggcgcttcaaaggaattctcaagatgcataggtcaatcagaggaagagcccgaggggagcagagatttcg
ccgcgagggaatcaatcactgtgccttgaagcaagcacttactgaaactaccgatgcgaccgcgcaagctgcaatctcgtaagctcg
tcgtcgagccgtatgaggtggaggccgatttcgcgaacctggactttccagaggaacttggcgacaatctttggcaggaggatgtgact
cgctcgagatggacttcgcgaaggggacgtcgaagcaaacgatacgcttgaaagcttcttcagcatatgccaggagcccaacgtgcttga
agtcgatttcttgtctcagtacctgggattggagacagagagcgtgcacgactgggttcaggggcaagcagaagaggggtcaagcatctct
ttgttacaagaaagctgccccaggggcaacagcgttcgcatgttggaccaccgcatgcgggcaaagtctgcacgacattttccctccgcgt
ctgagcatttctagataa]

```

protein sequence =

```

[MDNNEIIRLLALQDDYQPTAEAPEENSFCAAKEKPQDFFPPTVDDPDDNGTSPAGRIFEPFNKNRLEWEAFDRSKSCENDQAIADKLKE
LNSDVEESSGFNFNNSMVDSPGKTMSMNRTFLFQEGVDPRFQRLDDLTMILYRMEEGLANNMLTMINDGASKEFSRCIGQSEEEPEGQRF

```



RREAINHCALKQALTETTDATARSCNLVKLVVEPYEVEADFANLDFPEELGDNLWQEDVTRSRWTSRRDVEANDTLESFFSICQEPNVL  
EVDFLSQYLGLETESVHDWFRGKQKRVKHLFVTRKLPQGNSVRMLDHRMRQSLHDIFPPRLSISR]

end gene g3

###